2017

The Development of the Therapy Process Observational Coding System - In-Session Involvement

Emily J. Wheat
Virgina Commonwealth University

THE DEVELOPMENT OF THE THERAPY PROCESS OBSERVATIONAL CODING SYSTEM – IN-SESSION INVOLVEMENT

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

by

EMILY BUTT WHEAT
Bachelor of Arts, University of South Alabama, 2006
Masters of Arts, University of Richmond, 2009

Dissertation and Academic Advisor: Bryce D. McLeod, Ph.D.
Associate Professor, Department of Psychology

Virginia Commonwealth University
Richmond, Virginia
May, 2017

Acknowledgement

I would like to dedicate this dissertation to my husband, John Butt, and my parents, John and Patricia Wheat, who provided unconditional support and encouragement during my graduate studies. I would also like to acknowledge the contributions of my graduate advisor, Bryce McLeod, and the members of the TIMS lab who spent countless hours reviewing therapy sessions: I am grateful for your time and efforts.

Table of Contents

List of Tables

Page

List of Figures

Page

Abstract

THE DEVELOPMENT OF THE THERAPY PROCESS OBSERVATIONAL CODING SYSTEM – IN-SESSION INVOLVEMENT

By Emily Butt Wheat, Ph.D.

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, May, 2017

Dissertation and Academic Advisor: Bryce D. McLeod, Ph.D.
Associate Professor, Department of Psychology

In-session client involvement (i.e., participation in in-session therapeutic tasks) is hypothesized to be a necessary component of youth therapy and associated with positive outcomes. However, research on in-session client involvement has been slowed by definitional problems. At present, the field has not yet adopted a single definition of client involvement that is applicable across different theoretical orientations, which has impacted the measurement of this construct. To remedy this problem, the field needs to adopt a definition of in-session client involvement that includes important components (i.e., behavioral, affective, and cognitive) of this construct that applies across different theoretical orientations and use this definition to guide instrument development. The current study reports on the development and initial psychometric assessment

of the Therapy Process Observational Coding Scale – Involvement (TPOCS-I), an observational measure designed to capture in-session involvement for youth therapy. Treatment sessions (*N* = 895) were drawn from (a) 55 youth (ages 7-13 years; *M* = 9.89, *SD* = 1.71; 51.5% Caucasian; 58.8% male) who received standard cognitive-behavioral therapy, modular therapy, or usual care for youth anxiety; and (b) 51 youth (ages 7-14; *M* = 10.35, *SD* = 1.89; 86.3% Caucasian; 60.8% male) receiving standard cognitive behavioral therapy for youth anxiety. Sessions were independently scored by seven coders using observational instruments designed to assess involvement, alliance, therapist competence, and therapist interventions. Interrater reliability – intraclass correlation coefficients (2,2)—for the item scores averaged 0.73 (*SD* = 0.08) and 0.82 (*SD* = 0.08) for the Kendall and Child STEPS samples, respectively. The TPOCS-I scale and subscale (Behavioral, Affective, Cognitive, Positive, Negative) scores failed to demonstrate discriminant validity from the alliance. The use of two subscale configurations (i.e., Behavioral, Affective, Cognitive; and Positive, Negative) was not supported. These findings are discussed and future directions including measure development in a sample of youth with diverse diagnoses and the use of different perspectives in the measurement of in-session involvement.

Chapter One

Introduction

Statement of the Purpose

Client involvement, defined as a client's participation in in-session activities during the course of therapy, is considered critical to client outcomes in therapy (Karver, Handelsman, Fields, & Bickman, 2005; Tryon & Winograd, 2011). Conceptual models focused on how therapy works consider client involvement to be a critical ingredient of successful therapy (Doss, 2004; Hill, 2005). Indeed, client involvement in therapeutic activities, or the interventions of therapy such as exposure, cognitive restructuring, or problem-solving, is considered to be a necessary prerequisite for promoting positive outcomes across different types of therapy (Bohart & Tallman, 1999; Drieschner, Lammers, & van der Stark, 2004; Orlinsky, Grawe, & Parks, 1994). It therefore represents an important variable within therapy, as it is unlikely that clients will improve without participating in therapeutic interventions.[1]

The available research suggests that client in-session involvement[2] in therapeutic activities is linked to positive outcomes. In a review of the youth therapy process literature, Karver and colleagues (2006) found the average r-type effect size between client involvement and outcome was $r = .27$. This effect size estimate was based on 10 studies that had examined the client involvement-outcome association. More recent research provides further support for the

[1] In the current study, therapeutic interventions were measured using the Therapy Process Observational Coding System for Child Psychotherapy – Revised Strategies Scale (McLeod, Smith, Southam-Gerow, Weisz, & Kendall, 2015). Interventions assessed by this measure include theory-specific (e.g., cognitive interventions) and general interventions that are not orientation specific (e.g., encouraging affect; rehearsal; McLeod et al., 2015).

[2] Broadly, client involvement within the therapy process might include a number of activities both within session (e.g., in-session involvement; session attendance) and outside of session (e.g., homework completion). The definitely of client involvement for the current study is narrowed to a client's in-session involvement. Although this narrow definition may limit the applicability of findings to a specific subset of client behaviors, it provides the laser-like focus that is often necessary when detangling constructs that are closely related.

association between client and outcome (Chu & Kendall, 2004; Hudson et al., 2014). The limited evidence therefore suggests that client involvement may be linked to outcomes in youth therapy.

Despite the potential importance of in-session client involvement, relatively little research has been conducted in this area. To date, only 14 studies have investigated the involvement-outcome association in youth therapy (Chu & Kendall, 2004; Hudson et al., 2014; Karver et al., 2008; Karver et al., 2006; Lindheim & Kolko, 2010). Though these studies suggest that client involvement may be associated with outcomes, it is difficult to draw clear conclusions from the existing literature due to the limited number of studies.

Beyond the limited number of studies, another important problem facing the field is definitional as the field has not settled on a unified definition of client involvement (Morris, Fitzpatrick, & Renaud, 2014). Each study has used different definitions (e.g., energy spent in therapy; engagement in therapeutic tasks; homework completion; in-session behaviors), which presents a problem when attempting to compare and synthesize research related to client involvement (Martinez, Lewis, & Weiner, 2014). To address this problem, the field needs to adopt a common definition of this process variable.

Current measures of client involvement reflect the lack of an overarching definition of involvement. Involvement is considered by some to be a multi-faceted variable that consists of behavioral, affective, and cognitive components (Morris et al., 2014). Most existing measures emphasize involvement in behaviorally-focused activities (e.g., actively participated in interaction; Karver et al., 2008), but fail to have an equal emphasis on other activities of an affective or cognitive nature (e.g., identifying and correcting cognitive distortions; participating in exposures; identifying and understanding the effect of emotions). In addition, only two measures assess client involvement in youth therapy as a primary focus (Chu & Kendall, 2004;

Karver et al., 2008). Moreover, the distinction between these involvement measures and related constructs (i.e., the alliance) is poorly defined and this may lead to confusion when trying to tease apart individual effects and associations of these variables on each other and outcome.

To address these problems some have suggested the field needs a definition of client involvement that can be applied across different types of therapies (i.e., a pantheoretical definition) that consists of behavioral, affective, and cognitive components (Morris et al., 2014). The goal of this dissertation is to take an initial step toward developing an observational measure of client involvement that is appropriate for assessing the in-session participation of youth across a variety of therapies. This will be accomplished by using a definition of client involvement that includes components related to behavioral, affective, cognitive, positive, and negative involvement, and by differentiating client involvement from other related variables (i.e., alliance, therapy engagement, homework completion, and attendance).

To achieve this goal the current study proposes a pantheoretical definition of client involvement based on an examination of theory and historical perspectives. Afterwards, the important components of measure development when developing a measure for client involvement are described and existing client involvement measures are reviewed. Next the development of a measure of in-session client involvement using the proposed definition is described. This measure includes behavioral, affective, and cognitive components, and was used to code therapy sessions of a sample of youth with anxiety in three different treatment conditions: (a) manualized cognitive behavioral therapy (CBT) for anxiety (i.e., Coping Cat), (b) modular CBT for anxiety (i.e., MATCH), and (c) usual care. The resulting data were used to evaluate two hypotheses. First, it was hypothesized that the proposed measure would demonstrate score reliability in the measurement of in-session involvement. It was also

hypothesized that the measure would demonstrate score validity. Pearson correlations of measures assessing theoretically related constructs (i.e., involvement) were used to examine the extent to which scores demonstrate convergent validity. It was expected that measures of alliance would overlap the most with the proposed measure, followed by client centered therapeutic interventions. In addition, it was expected that behavioral interventions would correlate most strongly with behavioral involvement, affective interventions would correlate strongly with affective involvement, and cognitive interventions will correlate strongly with cognitive involvement. Finally, it was expected that involvement would vary across different clients, but not as a function of treatment condition, time, or therapist.

The proposed sample was a good fit for testing these hypotheses. First, CBT for youth anxiety includes behavioral, affective, and cognitive-focused tasks. This allowed expert judges a wide range of opportunities to observe a variety of different in-session involvement behaviors. In addition, the three different treatment conditions and availability of multiple sessions for each client across the course of treatment provided opportunities for the assessment of variance in involvement across condition and over time.

Chapter Two

Literature Review

Few studies have assessed the association between client involvement and outcome. Those that have examined this association use inconsistent definitions of client involvement, and this is likely because the definition of involvement varies within youth therapy. In addition, the measurement tools and methodology used when assessing client involvement use different definitions of involvement. The current section provides a review of these topics and an overview of the current state of issues related to the conceptualization and measurement of client involvement in youth therapy.

Client Involvement and Outcome Association

Client involvement is an important and necessary component of successful therapy and it has been posited that involvement is a prerequisite for client improvements. There is evidence of an involvement-outcome association in therapy for both adult and youth populations (Chu & Kendall, 2004; Gorin, 1993; Tobon et al., 2011; Tolan et al., 2002). However, the client involvement evidence base does not provide a clear picture of the role this variable may play in leading to the reduction of psychological distress. Specific definitions of client involvement vary and this leads to inconsistencies in the measurement of this variable. This presents a problem when attempts are made to draw conclusions from the client involvement literature.

Current research suggests that client involvement in therapeutic activities is linked to positive outcomes. In a review of the youth therapy process literature, Karver and colleagues (2006) found the average r-type effect size between client involvement and outcome was $r = .27$ based on 10 studies that examined the client involvement-outcome association. Limitations of this literature review include assuming an involvement-outcome association from concurrent

administration of process and outcome measures, the completion of process and outcome measures exclusively by the same informant, and inclusion of some studies that were not published in peer-reviewed journals. However, the primary limitation in these studies is definitional, for instance, some studies included in this review defined involvement as engagement in therapy and homework tasks and the use of inconsistent definitions of involvement in the review makes it difficult to ascertain how accurate this client-outcome association is.

Since Karver and colleagues' (2006) meta-analysis, a handful of studies have examined the client-outcome association in youth therapy (Chu & Kendall, 2004; Hudson et al., 2014; Karver et al., 2008; Lindheim & Kolko, 2010). The results of these studies support the findings of the Karver et al. (2006) meta-analysis and suggest client involvement is related to outcomes. These studies offer some methodological improvements over previous studies, such as multiple measurements of involvement over time (i.e., Hudson et al., 2014; Lindheim & Kolko, 2010) and comparison of involvement in different types of therapies (Karver et al., 2008). However, these studies also present with weaknesses in regards to definition. For instance, the measures used to assess involvement in these four studies primarily focus on behavioral aspects of involvement. This definitional bias towards behavioral involvement makes it is difficult to draw clear conclusions from the existing literature.

In summary, there is support for an involvement-outcome association, but these findings are limited by the small number of studies and their methodological shortcomings. Across these studies, involvement has been defined (and therefore assessed) in different ways (e.g., behavioral participation in session activities; attentiveness), limiting our ability to synthesize the existing evidence base. Other methodological considerations such as the perspective of involvement

measurement (i.e., who rates client involvement), and manner in which involvement is measured over time often limit the results of involvement-outcome findings.

Inconsistent Definition of Client Involvement

As previously discussed, client involvement within youth therapy is inconsistently defined. In addition, this variable has often been associated with other related variables such as therapy engagement, attendance, and homework completion. This can lead to confusion about where the boundaries between these related variables begin and end. Until there is a consistent definition of client involvement, it will be difficult to build and synthesize research surrounding this variable. Definitional issues can exist in many forms. For instance, researchers may conduct a study that assesses client involvement, but use a different term such as "engagement" (e.g., Lindheim & Kolko, 2010) or, researchers might conduct a study that assesses a different related variable, and label this variable "client involvement" (e.g., Richards, Bowers, Lazicki, Krall, & Jacobs, 2008). These definitional problems hinder the cross-study comparisons and the accumulation of the science knowledge base (Martinez et al., 2014). Therefore, a consistent definition of client involvement is necessary to advance our knowledge of this important factor and its role in youth therapy.

Defining client involvement.

An appropriate definition of client involvement would not only distinguish it from other related variables of interest, but also be applicable across different types of youth therapy. In one conceptual review of client involvement, key components of involvement across CBT, humanistic-existential-experiential therapy, and psychodynamic therapy are examined (Morris et al., 2014). Although the role and presentation of client involvement is different between approaches, the authors discovered that three key tasks are present across therapeutic

interventions. These tasks are cognitive, affective, or behavioral in nature. The authors concluded that involvement is a pantheoretical construct and therefore requires a definition that applies across theoretical orientations.

For the current study, client involvement is defined as the degree to which a client participates behaviorally, affectively, and cognitively in in-session therapeutic activities. More participation is indicative of higher involvement whereas a lack of participation, or resistance, is viewed as lower involvement. *Client* is defined as the individual who is the target of the intervention. *In-session* indicates that the behavior must occur during the therapy session and not as part of a homework assigned between sessions. This also excludes the behaviors that occur prior to session such as attendance and therapy engagement behaviors. It is necessary for a client to engage in therapy and attend for therapy to take place; however, these are not included in the current definition of client involvement. Finally, a child's participation (or lack thereof) is assessed only in regards to therapeutic activities; when no therapeutic activities are present, client involvement is not assessed.

Components of client involvement.

As noted by Morris and colleagues (2014), clients are provided with opportunities to participate in behavioral, affective, and cognitive tasks across therapeutic approaches. In the following sections, each component is defined.

Behavioral involvement. Behavioral involvement is defined as behaviorally participating in the tasks of therapy (Morris et al., 2014). This includes learning and practicing new skills such as relaxation techniques or roleplaying. Often clients learn about new behaviors and skills within the therapy environment including a rationale for how and why these tactics might be important in reducing distress. Clients might learn to problem solve by describing their problems, define

how these problems can be addressed, and evaluate how to approach this problem in a manner that is in line with their goals (Nichols, 2006, 2007). Additionally, they might identify ways to change their situation and the problem in session (De Jong & Berg, 2008). Afterwards, clients may practice within session. In youth therapy, youth might learn how to identify and address a problem, interact with peers in social situations, manage anger, respond to difficult situations in an assertive manner, or decrease distress through relaxation. Clients who are highly involved in behavioral therapeutic tasks are those who actively participate, and are attentively engaged during therapeutic skill presentation and rehearsal. Less involved clients might refuse to participate in behavioral tasks, attend to stimuli other than therapeutic material, or refuse to make choices when presented with options regarding therapeutic activities (Beutler, Moliero, & Talebi, 2002).

Affective involvement. Affective involvement is defined as the degree to which emotions are experienced, labeled, and processed (Morris et al., 2014). This includes any activity designed to help clients connect with emotions including labeling and accepting emotions, mindfulness activities, body scanning, and exposure. Client affective experiences are often discussed in the early phases of therapy, and in some approaches, may be discussed through the course of treatment (Glasser, 1992; Rogers, 1987). In youth therapy, youth might learn about emotions including how to label them and use body cues to identify what emotion they are feeling, describe and process emotional reactions to meaningful events, become more aware of how they are feeling, or practice approaching difficult emotion-charged experiences in the safety of the therapy environment (e.g., exposure). Youth who are more involved in affective therapeutic activities participate willingly and with enthusiasm, and describe and elaborate on their

emotional experience. Less involved youth might avoid experiencing the emotions elicited in affect tasks.

Cognitive involvement. Cognitive involvement is defined as exploring and commenting on thought processes (Morris et al., 2014). Another important aspect of client involvement that spans therapeutic approaches is the idea of self-acceptance, insight, and changing maladaptive ways of thinking. Early in therapy, clients are often asked to share their unique experience including exploring thoughts about themselves, others, and their circumstances (Carlson, Watts, & Maniacci, 2006; Yalom, 2003). A focus on one's cognitions and perspectives might remain a main focus of the therapy experience (Ellis, 1994, 1996). In youth therapy, youth might learn about distorted versus logical styles of thinking and how to differentiate between the two, practice identifying distorted thoughts, and practice altering distorted thoughts into more adaptive ones. Youth who are more engaged are those who share their thoughts by self-disclosing and initiating new topics of discussion, and eagerly participate in tasks related to identifying and adapting distorted cognitions. Less involved youth might withdraw or refuse to respond to the therapist.

Terms often associated with client involvement.

In defining client involvement it is important to distinguish it from other constructs that are similar. Client involvement has often been associated with a number of related variables (i.e., therapy engagement, attendance, and homework completion). The proposed definition of client involvement (i.e., the degree to which a client participates behaviorally, affectively, and cognitively in in-session therapeutic activities) is distinct from these variables, as discussed below.

Therapy engagement. Therapy engagement is defined as the initial therapy attendance and ongoing therapy efforts including attendance, homework completion, and in-session involvement (McKay & Bannon, 2004). This term describes a broad process, whereas client involvement is a specific term that refers to in-session client behavior. Although there is some overlap between these two terms, in that both client involvement and therapy engagement include in-session involvement, the current proposal does not focus on the other concepts included within the therapy engagement process (e.g., treatment attendance).

Attendance. Attendance refers to the number of sessions where the child, family, or other individuals in the child's life (e.g., teacher) are present (Garland, Haine-Schlagel, Accurso, Baker-Ericzén, & Brookman-Frazee, 2012). Attendance is necessary for in-session client involvement to occur, but is distinguished from involvement in that attendance is a matter of a client being physically present at a session. Attendance does not include the examination of the degree to which a client participates within a session.

Homework completion. Homework completion is defined as the degree to which a client completes tasks outside of session. These tasks are assigned by the therapist and examples include applying coping skills in a difficult situation, completing at-home exposures, and practicing deep breathing (Burns & Nolen-Hoeksema, 1991; De Araujo, Ito, & Marks, 1996; Edelman & Chambless, 1993; Kazantzis, Deane, & Ronan, 2000). Homework completion and homework adherence are terms used to describe the same concept (Clarke et al., 2013). Homework completion might be discussed in session as part of treatment activities. Homework completion can be distinguished from in-session involvement because the work required to complete homework occurs outside of session. A therapist may review or assign homework

during session and a client's participation during these moments is considered when rating the degree of in-session involvement, but any work outside of session is not.

The Role of Client Involvement in Youth Therapy: A Therapy Process Model

Not only is client involvement viewed as an important variable across therapeutic approaches, it is a variable that has demonstrated associations with other key variables of the therapy process. A conceptual model (Figure 1) depicts how client involvement is associated with other variables of interest (i.e., client and therapist pre-treatment characteristics, therapist interventions and competence, alliance, and outcome). This model is an adaptation from Fjermestad, McLeod, Tully, and Liber (2015).

In this section, a general overview of all components of the model is discussed, including definitions of each term. The discussion of each component will include a general definition and description of broad relationships this component is proposed to have with other therapy process variables, followed by a specific explanation of how these components are hypothesized to relate to client involvement.

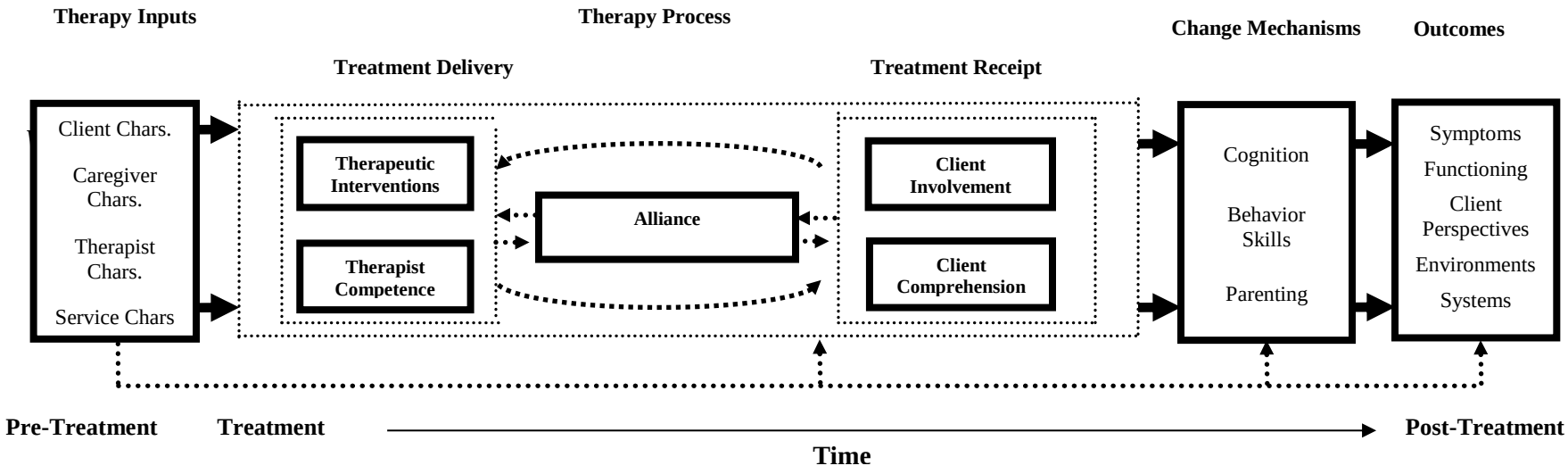

Figure 1. A model of therapy process for youth therapy; Chars = characteristics. *This figure is an adaptation of Fjermestad and colleagues (2015) model of therapy process and presents the important components that are present in youth therapy.*

Therapy inputs. Therapy inputs include youth characteristics and therapist characteristics and are proposed to influence or moderate the process and outcome of youth therapy. Inputs are factors that are present at the beginning of treatment. Youth characteristics might include sex, race and ethnicity, socioeconomic factors, age, diagnosis, and symptom severity (Chorpita, Daleiden, & Weisz, 2005). Therapist characteristics might include those related to training and theoretical orientation, demographics, or interpersonal style (Lambert, 2010; Shirk & McMakin, 2008). One would expect for client characteristics to be associated with subsequent therapy process factors, including client involvement.

In therapy, it is hypothesized that client and therapist pretreatment characteristics influence subsequent involvement and there is some evidence to support this claim. Parent demographic attributes, stressors, and child behavior problems were related to the quality of parent participation in parent management training (Nix, Bierman, & McMahon, 2009). Similarly, client symptom severity for adult participants was a predictor of client involvement in behavioral in-session exercises (Edelman & Chambless, 1993). In addition, there is some evidence to support an association between client race and initial client involvement, and an association between parental mood disorder diagnosis and change in client involvement over time (Becker et al., 2014). But, it is difficult to draw firm conclusions about involvement from the findings of Becker and colleagues (2014) because involvement was assessed as part of therapy engagement. Although little research has examined the impact of client and therapist pretreatment factors on client involvement in youth therapy, it is hypothesized that identifying and attending to the specific pretreatment client characteristics likely to impact involvement may improve youth in-session participation and afterwards lead to improved client outcomes (Chu, Suveg, Creed, & Kendall, 2010).

Therapy process factors. Therapy process factors include elements related to therapist delivery of therapeutic interventions (i.e., type of intervention delivered and therapist competence in intervention delivery), the alliance, and factors related to treatment receipt (client involvement and client comprehension; Burgio et al., 2001; Fjermestad et al., 2015).

Treatment delivery. Treatment delivery includes therapist actions and their ability to present treatment as intended (Lichstein, Riedel, & Grieve, 1994). Therapist actions, or therapeutic interventions are defined as techniques used in conjunction with a particular theoretical orientation aimed at producing change in a client's presenting problem. Some propose that specific therapeutic interventions are associated with symptom reduction (McLeod & Weisz, 2005) while others propose that interventions combined with patient, therapist, contextual, and relational factors lead to improvements in outcome (Beutler, 2014). It is also hypothesized that specific interventions lead to increases in client involvement, whereas other interventions are expected to hinder client involvement (Jungbluth & Shirk, 2009). There is some evidence supporting this claim. For instance, subsequent client involvement was predicted by therapist behaviors of exploring motivation and providing structure (Jungbluth & Shirk, 2009). In addition, when therapist relationship-building behaviors increase or decrease, client involvement and alliance have been found to follow suit by either increasing or decreasing along with these therapist behaviors (Hudson et al., 2014). Given these findings, one would expect client-centered interventions (e.g., exploring client motivation, providing structure to treatment sessions, and incorporating relationship-building behaviors) to exhibit a positive relationship with in-session involvement.

Therapist competence is a term that refers to the quality of treatment delivery (Southam-Gerow & McLeod, 2013) and can refer to competence in delivery of specific interventions as

part of a larger therapy approach (e.g., exposure in CBT; Kendall, 1994) or competence in general interventions believed to be important across different therapeutic approaches (e.g., alliance-building behaviors; Diamond, Hogue, Liddle, & Dakof, 1999). It is hypothesized that competence will lead to positive client outcomes by maximizing client involvement and the alliance. This hypothesis was examined by Chu and Kendall (2009) with a sample of youth receiving CBT for anxiety. Correlational analyses supported this model, but results were not significant when formally tested using a moderated mediation model. This group conducted a second study with a larger sample of youth receiving CBT for anxiety to examine this model (Hudson et al., 2014). Results indicated that increases in therapist flexibility did not lead to increases in involvement or alliance. At this time, it is unclear how therapist competence in treatment delivery might influence client involvement in youth treatment. Empirical research examining the association between involvement and competence within the adult literature is also sparse. However, there is some evidence to suggest that therapist competence and client involvement are positively associated and that therapist competence is perceived to be high when client involvement in treatment is also high (Hill et al., 2006).

The alliance. The alliance is defined as the affective and collaborative aspects of the client-therapist relationship (McLeod, 2011). Current evidence indicates that client characteristics may play a role in alliance formation (e.g., Flicker, Turner, Waldron, Brody, & Ozechowski, 2008) and specific therapist interventions facilitate the therapist-child alliance (e.g., Diamond, Liddle, Hogue, & Dakof, 1999). In addition, meta-analyses examining the strength of the alliance-outcome association have estimated small r-type effect sizes (r = 0.14 to 0.21; Karver et al., 2006; McLeod, 2011).

It has been hypothesized that a strong alliance leads to increases in client involvement and some evidence supports this claim (McLeod, Islam, Chui, Smith, Chu, & Wood, 2013). Youth-therapist alliance predicted subsequent client involvement for adolescents receiving CBT and non-directive supportive treatment for depressive symptoms after attempted suicide (Karver et al., 2008). Likewise, positive changes in alliance predicted subsequent involvement for youth completing child-focused and family-focused CBT for anxiety (McLeod et al., 2013). Alternatively, Hudson and colleagues (2014) obtained observational measures of the alliance and involvement from eight time points of therapy for youth participating in manual-based CBT for anxiety. Results indicated that alliance and involvement early in treatment were highly correlated, and that increases in alliance were associated with increases in involvement over time (Hudson et al., 2014). Based on these findings, client involvement and alliance appear to share a complex and intertwined relationship; and it is expected that measures of alliance and client involvement will be associated in youth therapy. However, it is also expected that client involvement will perform as a distinct variable from the alliance, as these are hypothesized to be two distinct variables.

Treatment receipt. Treatment receipt is a term used to describe client in-session behaviors including client involvement in in-session activities and client understanding of therapeutic concepts (Bellg et al., 2004; Burgio et al., 2001; Fjermestad et al., 2015). Treatment receipt relates to a patient's ability to demonstrate understanding and application of skills during the treatment session (Bellg et al., 2004). Thus, components of treatment receipt include client comprehension and client involvement. Client involvement is the primary focus of this proposal. As previously described, involvement is characterized as in-session involvement or participation in therapeutic tasks.

Change mechanisms. It has been hypothesized that client involvement in therapeutic tasks facilitates the change process (e.g., leading to decreases in cognitive distortions) and predicts improvements in outcome (Shirk, Crisostomo, Jungbluth, & Gudmundson, 2013). Change mechanisms are influenced by therapy process variables and lead to overall change in outcome (e.g., behavioral activation or decrease in cognitive distortions; Doss, 2004). Only one youth therapy study to date has examined the link between involvement, change mechanisms, and client outcome (Shirk, et al., 2013). Shirk and colleagues examined client involvement in cognitive tasks for youth receiving CBT for depression and the relation between involvement, change in cognitive distortions, and symptoms of depression. Although change in cognitive distortions predicted change in depressive symptomology, involvement (i.e., in-session participation and out of session homework completion) did not predict either of these variables. This study is the first to examine the relation between client involvement, change mechanisms, and client outcome in youth therapy. Although findings did not support the proposed hypothesis, methodological flaws limited the clarity of results. These flaws include the small sample size (N = 83), self-report assessment of cognitive distortions and outcome, and the limited definition and measure of client involvement only during skill presentation and practice of each session.

Outcomes. Client involvement is hypothesized to lead to improvements in client outcomes (i.e., symptom reduction and global functioning). As previously stated, Karver and colleagues (2006) found the involvement to outcome effect to be r = .27, but these findings were

based on studies with inconsistent definitions of involvement. Since Karver's (2006) review, Chu and Kendall (2004) examined the client involvement – outcome association and found that client involvement in manualized CBT for youth anxiety was predictive of treatment gains. Similarly, Hudson and colleagues (2014) found that involvement was associated with positive treatment outcomes for youth receiving manual-based CBT for anxiety. In another study examining CBT and nondirective supportive therapy for adolescent depression, the authors failed to find an association between involvement and outcome (Karver et al., 2008).

Client involvement is hypothesized to lead to improved outcomes (Doss, 2004; Hill, 2005). For instance, Doss (2004) suggests that the combination of therapist and client behaviors in session leads to improvements in overall outcome via mechanisms of change (i.e., adoption of coping thoughts). Client involvement is viewed as a necessary component for symptom reduction and global improvement.

In summary, in-session involvement is an important component of the youth therapy. The current conceptual model is used to generate hypotheses about how youth therapy works. One hypothesis is that youth in-session involvement is high when therapists administer client-centered interventions and when alliance is also high. This has been supported by previous research, therefore, it is expected that a similar patterns of results will be attained using a pantheoretical measure of in-session involvement.

Chapter Three

Measurement of Client Involvement in Youth Therapy

In-session involvement in youth therapy is necessary for therapeutic change to take place, regardless of the therapeutic intervention received or context in which therapy is provided. It is essential that measures of involvement use a pantheoretical definition to capture all facets (i.e., behavioral, affective, cognitive) of this important variable. Current measures of involvement are flawed because they omit critical components of involvement and are limited by the methods used to assess involvement. The current section reviews methodological components necessary to consider for the development and examination of an involvement measure including: (a) definition of target variable; (b) perspective of measurement; (c) target of measurement; (d) unit of measurement; and (e) type of measurement.

Definition of Target Variable

The definition of the target variable is perhaps the most critical variable of interest to measure developers (DeVellis, 2003). The definition of the target variable within a measure of involvement is important to consider because the term "involvement" could be used to describe a number of important aspects of the therapy process. As previously discussed, current measures of involvement were developed using different definitions and this has led to confusion about what is actually being measured (e.g., involvement or attendance).

Perspective of Measurement

Perspective of measurement concerns the viewpoint of the individual rating the measure. Each participant in the therapy session may have a slightly different view of a child's involvement. Therapy process measures can be rated from different perspectives such as the client (i.e., child and/or parents), therapist, supervisor, or outside observer. Child reports of

assessment are valuable in that youth have the potential to be the most accurate, as they would not need to make inferences about their own internal states and actions in regards to their participation in therapeutic activities. However, youth may have difficulty completing self-report measures accurately due to limitations in their cognitive developmental stage, especially when asked about affective tasks (Chambers & Johnston, 2002). For this reason, parents and therapists may be better equipped to make ratings of child involvement. However, parents and therapists will have to make inferences when making ratings about a child's internal state.

There is a paucity of research comparing measurement perspectives of youth in-session involvement. Currently, no research has examined the differences in reports from youth, therapists, and parents regarding child's level of involvement in therapy, so there are no data to imply which perspective is best. However, self-report is considered to include bias (Herbert, Ma, Ebbeling, Matthews, & Ockene, 2001). Further, parent- and therapist-report may also include bias related to the rater's ability to attend to various dimensions of involvement, or value placed on involvement dimensions outside of a specific orientation or treatment approach. These factors may lead to over- or under-estimation of a child's level of involvement. Ratings by trained observers may be unbound by these limitations and provide more information than self-report measures. For instance, observers may receive training across orientations and techniques, detect behaviors that may otherwise be overlooked, and make ratings without bias caused by personal interactions with the individual being observed. Finally, depending on stimulus materials, observers have the opportunity to pause, rewind, and re-watch sessions to maximize on data collection during busy or fast-moving sections of a therapy session.

Past measures of relational factors (i.e., alliance and involvement) have been rated by outside observers (e.g., Chu & Kendall, 2004; Karver et al., 2008), the therapist (Adelman,

Kaiser-Boyd, & Taylor, 1984), or the client (e.g., Horvath & Greenberg, 1989). Although time and resource intensive, the use of outside observers is often preferred to address concerns about informant discrepancies (e.g., discrepancies between parent and child perspectives of involvement) and capture behavioral processes (Bakeman & Quera, 2012).

Target of Measurement

Target refers to the individual (e.g., client, therapist, supervisor, combination) who is the center of attention for measurement ratings. This aspect of measurement is important when assessing in-session involvement because some sessions include multiple participants. Previous examination of client involvement has targeted youth (e.g., Chu & Kendall, 2004; Tolan et al., 2002) or parents (e.g., Noser & Bickman, 2000; Littell, 2001). Because both parent and youth involvement have been linked to client outcomes, measurement of one or both of these targets is key when assessing involvement (Karver et al., 2006). Oftentimes, manualized therapy sessions are targeted at children who experience mental health problems and contain in-session activities and tasks within an individual therapy format (Barkley, 1997; Chorpita & Weisz, 2005; Kendall, 1994; Weisz, Moore, Southam-Gerow, Weersing, Valeri, & McCarty, 2005). Therefore, it is important that a child's behaviors be assessed when examining involvement of child-focused therapy. Examining parent involvement may also be important for certain therapy modalities, as parents might be asked to participate in their child's therapy for certain presenting problems (e.g., Nix, Bierman, McMahon, & The Conduct Problems Prevention Research Group, 2009).

Unit of Measurement

Another important aspect of therapy process measures is the unit of measurement and this refers to the sample of behavior used to generate ratings. This aspect of measurement is important when assessing in-session involvement because a participant's involvement is likely to

change over time. Smaller units of measurement (e.g., utterances, speaking turns) focus on how a variable might change over the course of a session, whereas larger units (e.g., therapy sessions) have a more global focus and examine how a variable changes over the course of therapy (McLeod, Islam, & Wheat, 2013). When examining process measures, it is recommended that a more global approach be used to allow examination of process-outcome relation (Hogue, Liddle, & Rowe, 1996) or process-process relation (e.g., therapist behaviors and alliance formation; Creed & Kendall, 2005). This allows for an average estimate of client involvement across a session, instead of the examination of client involvement in response to specific interventions within a session. This approach is consistent with the conceptualization of client involvement as a treatment variable that is important regardless of therapy orientation.

Type of Measurement

The method of item response is important to consider because some methods may be better suited for the assessment of involvement than others. Interval scales are frequently used when assessing client involvement (e.g., Chu & Kendall, 2004; Karver et al., 2008). One important consideration when choosing interval measures is the number of responses: too many responses may be too difficult for raters to discriminate among and reach reliability, whereas too few responses may not be sensitive enough (DeVellis, 2003). Another key consideration when using interval scales is whether the scale should contain an even or an odd number of responses. An odd number of responses allows raters to choose a neutral item, whereas an even number of responses forces raters to make a commitment, in some degree, to one extreme or the other (DeVellis, 2003). In relational measures, choosing an even number of responses is often preferred (e.g., Chu & Kendall, 2004; McLeod & Weisz, 2005) as this forces raters to make a choice about the degree of client involvement or alliance within session.

Involvement-focused Measures

At this time, only two measures of client involvement exist in the youth therapy literature, the Child Involvement Rating Scale (CIRS; Chu & Kendall, 1999) and the Shirk Collaboration Coding Scale (SCCS; Karver et al., 2008). The CIRS and SCCS have been used to assess involvement in therapy for youth receiving therapy for symptoms of anxiety and depression, respectively (e.g., Chu & Kendall, 2004; Karver et al., 2008). The existing psychometric data surrounding these measures are reviewed with attention paid to the methodological factors previously discussed (i.e., definition of target variable, perspective, target, unit, and type of measurement).

Child Involvement Rating Scale. The Child Involvement Rating Scale is a youth-focused in-session involvement measure (CIRS; Chu & Kendall, 1999). The CIRS was originally developed as a 10-item observational measure to examine the positive (6 items, e.g., "Does the child make suggestions to change the task suggested by the therapist?") and negative (4 items, e.g., "Is the child oppositional to therapist suggestions and treatment activities?") aspects of involvement. Involvement was defined as active behavioral participation and openness to therapy and only in-session behaviors were assessed via use of audio recordings (Chu & Kendall, 2004). Initial psychometric properties were examined with a sample of 59 youth (ages 8-14 years old) with principal anxiety diagnoses (Chu & Kendall, 2009). Youth were recruited, treated, and assessed as part of three clinical trials (i.e., Flannery-Schroeder & Kendall, 2000; Kendall, 1994; Kendall et al., 1997) that took place within a university-based outpatient clinic. Youth received a 16-session manual-based CBT program and coders rated two 10-minute segments of two early therapy sessions and two later sessions from the first half of therapy (i.e., Sessions 2 to 5 and session 6 to 10, respectively). Ratings were made on a 6-point scale (i.e., 0 = *not at all present*; 5

= *a great deal present*) based on the frequency and significance of an event. Items with low variability and low correlations to remaining items were dropped, resulting in a six-item scale. Overall scores were computed by reverse-scoring the two negative items and summing these with the four positive items.

Psychometric analysis of the CIRS revealed adequate internal consistency (Cronbach's $\alpha = 0.73$) and inter-rater reliability as measured by intraclass correlations (ICC = 0.61) was good. In addition the CIRS demonstrated predictive validity; higher levels of involvement during sessions 6 to 10 was associated with improved outcome (i.e., absence of child's principal anxiety diagnosis or improvement in impairment ratings at post-therapy; Chu & Kendall, 2004).

Psychometric properties of the CIRS scores were also examined in a sample of 34 youths (age 8-12 years old) receiving group CBT (GCBT) for a principal anxiety disorder diagnosis (Tobon et al., 2011). Coders rated all 12 sessions of manualized GCBT and ratings were made based on watching the entire sessions live or on video. Mean of inter-item correlations was adequate (expected range is .15 to .50; Clark & Watson, 2003) at $r = .48$ (range =.00-.97). Internal consistency ($\alpha = .89$) was strong. Inter-rater reliability was good to excellent with average ICCs at .74 (range = .63-.90).

The involvement variable was examined across cases. Chu and Kendall (2004) found the CIRS revealed an overall negative involvement shift over time. Toban et al (2011) using the same involvement scoring, found early involvement (average score from sessions 1 to 6) predicted post-therapy anxiety symptoms. When participants from Chu and Kendall's (2004) individual CBT study were included with the Toban and colleagues' sample, early involvement predicted post-therapy anxiety symptoms when controlling for pre-therapy anxiety symptom ratings and late involvement (Tobon et al., 2011). Unlike the findings in Chu and Kendall's

(2004) study, the mean overall involvement scores in Toban and colleague's (2011) study were relatively stable over time.

Shirk Collaboration Coding Scale. The Shirk Collaboration Coding Scale (SCCS; Karver et al., 2008) is a youth-focused measure of in-session involvement. Client involvement was defined as cooperating with, being involved in, making suggestions about, and/or completing therapeutic tasks, including homework (Karver et al., 2008). Only in-session behaviors were evaluated. The SCCS is comprised of five items rated by observers and developed from the Patient Participation Scale of the Vanderbilt Psychotherapy Process Scales (VPPS; O'Malley, Suh, & Strupp,, 1983). Items are rated on a 5-point scale (1 = *not at all* to 5 = *a great deal*). The SCCS was initially utilized to examine the association between therapist engagement strategies, therapist-client alliance, and client participation. The psychometric properties of the SCCS scores were initially examined in a sample of 23 youth (ages 13-17) with depressive symptoms who had attempted suicide (Karver et al., 2008). Youth were recruited, treated, and assessed as part of a randomized controlled trial comparing CBT to nondirective supportive therapy (NST) in an outpatient treatment setting.

Coders listened to a 10-minute segment from the fourth session of therapy after the introduction of a therapeutic task (i.e., social problem solving or request for information about internal state). Coding integrity was maintained by resolving discrepancies through comparing coding notes and discussion and resulted in excellent reliability (ICC = 0.95; Karver et al., 2008). Client involvement at session four showed a trending association with positive change in depressive symptoms from baseline to post-therapy for the CBT condition ($r = 0.56$; $p = 0.095$; Karver et al., 2008). A strong and significant association was found between alliance and involvement ($r = 0.76$).

Summary of involvement measures. Overall, existing evidence suggests that scores on the CIRS and SCCS demonstrate predictive validity for youth outcomes in therapy for anxiety and depression, respectively. Studies using the CIRS provided global measures of involvement, whereas the SCCS was measured at one time point only, a strategy authors have suggested is less appropriate when attempting to examine process-outcome or process-process associations (Karver et al., 2008).

In regards to the measurement of the construct of involvement, the CIRS and SCCS focus primarily on behavioral involvement within session and this definition may limit the utility of these measures in non-directive therapies. In addition, studies examining the CIRS and SCCS have lacked examination of discriminant validity. For instance, it is unclear the degree to which these measures are distinct from related constructs, such as the alliance. Indeed, the SCCS was found to be highly correlated with measures of alliance and this association has not been formally tested with the CIRS. In addition, the SCCS included homework completion within the definition of involvement instead of separating these distinct involvement behaviors as separate constructs.

Last, there are methodological limitations that may weaken the generalizability of the results from these two observational measures. Both the CIRS and SCCS were rated using audio recordings as stimuli and ratings were based on listening to a portion of the session. Currently, there is no evidence related to the ideal stimuli and length of observations for in-session involvement, but there is potential to miss important data when observations are made from small portions of sessions.

In summary, both the CIRS and SCCS are youth-focused observational measures of in-session involvement. The strengths of the existing measures are their use of expert coders and

adequate initial psychometric properties (e.g., reliability). However, there are several limitations including their strong focus on behavioral involvement, and potential overlap with the alliance, a related but distinct variable.

Measures of therapy process

In-session client involvement is also assessed as part of larger therapy process assessment systems. Two of the most utilized measures for therapy process, also parent measures that were used in the development of the CIRS and SCCS, are the Child Psychotherapy Process Scales (CPPS; Estrada & Russell, 1999) and the Vanderbilt Psychotherapy Process Scales (VPPS; Gomes-Schwartz, 1978). Because of the pervasiveness of these measures in the involvement literature, and their seminal importance to the two existing measures of youth involvement, these measures are also reviewed here to fully cover the assessment of involvement in therapy.

Child Psychotherapy Process Scales. The Child Psychotherapy Process Scales (CPPS, Estrada & Russell, 1999) measure youth therapy process. This measure is comprised of 33 items rated by expert observers and is primarily focused on the relationship between the youth and therapist. A factor analysis of the CPPS revealed a factor (i.e., Child Therapeutic Work) related to behavioral in-session involvement and affective involvement and included items related to feelings, understanding, productivity, exploration, and engagement (Estrada & Russell, 1999). No formal definition of client involvement is provided. The wording of these items makes it difficult to determine if youth affective responses occur as a result of the therapeutic task or the client-therapist relationship (e.g., actively participated in the interaction; was noncompliant with the therapist and with the tasks of therapy). The psychometric properties of the CPPS were initially examined in a sample of 13 youth (aged 6-12) receiving long-term individual therapy for oppositional defiant disorder, attention deficit hyperactivity disorder, dysthymia, posttraumatic

stress disorder, or pervasive developmental disorder. Youth were recruited, treated, and assessed within a university-affiliated mental health training center following a broadly defined psychodynamic model including a combination of verbal and play therapy (Estrada & Russell, 1999). Objective coders made ratings of session transcripts on a 5-point scale assessing the extent to which each item was present. A total of 35 therapy transcripts were divided into verbal exchanges and the total of exchanges were divided into thirds for each client. CPPS scores were created by subtracting the average ratings across all three segments for a given client from the average of the positive items. Since its development, the CPPS has also been used in observational coding of audio- and video-taped sessions of youth participating in randomized controlled trials (i.e., Kendall, Hudson, Gosch, Flannery-Schroder, & Suveg, 2008; Kendall et al., 1997; Flannery-Schroder & Kendall, 2000; Kendall, 1994) for anxiety (Hudson et al., 2014).

Reliability as measured by agreement between coders was in the good range (ICC = 0.77). Internal consistency had a wide range between unacceptable to good (therapist items α = 0.49 - 0.88; youth items = 0.68 to 0.83). The subscale, Child Therapeutic Work, emerged as one factor when data were submitted to factor analysis. There is equivocal evidence about discriminant validity. The Child Therapeutic Work subscale was not significantly or strongly associated with the Therapist Therapeutic Relationship subscale (suggesting that there is no relationship between involvement- and alliance-focused subscales). However, there is also some evidence to suggest that the alliance-focused items might not discriminate between involvement as assessed by other measures. For example, when the CPSS was used in conjunction with the CIRS (previously described as an involvement measure for youth therapy), a factor analysis revealed that alliance related items on the CPPS were found to overlap with CIRS items (Hudson et al., 2014).

Vanderbilt Psychotherapy Process Scale. The Vanderbilt Psychotherapy Process Scale (VPPS; Gomez-Schwartz, 1978) was originally developed as an 84-item observational measure to assess significant aspects of therapy process for adult-focused therapy. Involvement was defined as the extent to which a client actively engaged in therapy interaction. Items are rated on a scale from 1 (*not at all*) to 5 (*a great deal*). This measure is predominantly focused on the examination of the client's and therapist's behavior and attitudes, and includes a client participation subscale with items focused on behavioral, affective, and cognitive indicators of involvement. Measurement intervals may vary, but the VPPS manual recommends that coders base ratings on observations of therapy sessions at least five (preferably 10-15) minutes in length (Suh, O'Malley, Strupp, & Johnson, 1989).

The VPPS has undergone at least three revisions aimed at improving the instrument. Across these revisions, the VPPS has demonstrated adequate levels of both internal consistency (0.96-0.82) and interrater reliability (0.94-0.79) for all subscales (Krupnick et al, 1996). The Patient Participation subscale of the VPPS has been associated with client outcomes (Gomes-Schwartz, 1978; O'Malley et al., 1983). For instance, of all of the VPPS process scales, the client participation scale was the strongest and most consistent predictor of client outcome post-treatment, regardless of perspective of outcome rating (e.g., therapist, client, or observer; Gomes-Schwartz, 1978). Participants in the Gomes-Schwartz (1978) study included college males who were attending therapy at a university counseling center for difficulties with anxiety, shyness, and interpersonal relations. O'Malley and colleagues (1983) found similar results when examining the involvement-outcome association in a sample of college males presenting as socially introverted, depressed, and obsessive-compulsive features. Observer-rated involvement predicted overall improvement and improvement of target complaints across three perspectives

(e.g., clinical interviewer, therapist, and client; O'Malley et al., 1983). See Suh, O'Malley, Strupp, & Johnson (1989) for an in depth review of VPPS psychometrics.

Summary of therapy process measures. There are numerous strengths to the measures discussed including adaptation to youth populations, strong psychometric properties, and predictive utility. Overall, existing evidence reveals that the VPPS is the most widely used and validated of the measures reviewed. This adult-focused measure has been adapted for use with youth populations, and initial evaluation suggests these adaptations (e.g., Vanderbilt Therapeutic Alliance Scale) are useful in assessing youth therapy process. The VPPS has reliably demonstrated predictive validity for process-outcome association, including involvement-outcome association, in adult populations. The CPPS also demonstrates good internal consistency and reliability data.

Despite strengths, there are limitations in relying upon existing measures to assess youth involvement. First, limited construct definition and adult sample development hinder any accurate conclusions drawn from downward extensions of adult measures (e.g., VTAS). Limitations within each measure make these inadequate tools for assessing client involvement across different therapeutic approaches as they lack a pantheoretical definition. For example, the VPPS lacks clear assessment of client cognitive involvement. The only youth-focused measure, the CPPS, may fail to differentiate between the alliance and involvement, two related but distinct variables. These measures, therefore, are insufficient when attempting to examine all aspects of client involvement within session.

Alliance measures with involvement components

In addition to examining measures of involvement and general therapy process, measures of alliance were reviewed. Although alliance and involvement are two distinct variables, the

distinction between them is often unclear due to the manner in which they are defined and assessed. Measures of alliance for youth treatment are reviewed in an attempt to define areas of potential overlap. This information was used to decrease definitional ambiguity in the development the measure of in-session involvement

Alliance Observation Rating Sheet. The Alliance Observation Rating Sheet (AOR; Karver et al., 2008) is a youth-focused observer rated 15-item measure developed for assessing client behaviors that contribute to the alliance, defined as the relational, emotional, and cognitive connection between youth and therapist. The psychometric properties of the AOR were initially examined in the same sample as the SCCS (i.e., 23 youth between the ages of 13 and 17 with depressive symptoms who had attempted suicide; Karver et al., 2008). Youth were part of a randomized controlled trial comparing CBT to NST. Coders watched the third session of therapy and made ratings of behaviors based on observation of three 20-minutes blocks. Ratings were made on a 5-point frequency of occurrence scale and session scores were created by summing ratings of the 15 items.

There was high agreement between independent coders (ICC = 0.84) and discrepancies were resolved by averaging coder ratings. Items were highly correlated with one another and internal consistency was high (Cronbach's $\alpha = 0.92$). Additionally, the AOR was significantly correlated with self-report ratings of the alliance (e.g., Working Alliance Inventory – Self Report), demonstrating construct validity.

Therapy Process Observational Coding System – Alliance Scale. The Therapy Process Observational Coding System for Child Psychotherapy – Alliance Scale (TPOCS-A; McLeod & Weisz, 2005) is a 9-item observational measure that examines the youth-therapist alliance. Alliance was defined as a construct with two dimensions: (a) the bond, or affective

aspects of the client-therapist relationship, and (b) the task, or client participation in the activities of therapy (McLeod & Weisz, 2005). The psychometric properties of the TPOCS-A were initially examined in a sample of 22 youth (ages 8-14) with principal anxiety or depressive disorders. Youth were recruited, treated, and assessed as part of the Youth Anxiety and Depression Study (YADS; Weisz et al., 2009) that took place within five outpatient community mental health clinics. Coders watched four randomly selected sessions (i.e., therapy was divided into thirds and one session from the beginning, two from the middle, and one from the end of therapy were selected). Items on the measure, initially 14-items, were rated on a 6-point scale (0 = *not at all*; 5 = *a great deal*). Youth received usual care (or UC) and coders rated entire therapy sessions.

The final 9-item TPOCS-A scale demonstrated adequate inter-rater reliability (ICC item $M = 0.59$) and acceptable internal consistency ($\alpha = .95$). The TPOCS-A also demonstrated adequate convergent validity when compared to the Therapeutic Alliance Scale for Children (TASC; Shirk & Saiz, 1992), a self-report measure, with correlations between observed and self-report alliance at .53 (McLeod & Weisz, 2005). In addition, the TPOCS-A has demonstrated predictive utility. For example, youth alliance was associated with youth outcomes for youth treated for internalizing disorders, even after controlling for confounding variables (McLeod & Weisz, 2005). The TPOCS-A has also been used to assess alliance in youth diagnosed with ADHD receiving friendship training (Lerner, Mikami, & McLeod, 2011), and youth participating in a randomized controlled trial comparing youth-focused and family-focused CBT for anxiety (Chiu, McLeod, Har, & Wood, 2009).

Vanderbilt Therapeutic Alliance Scales. The Vanderbilt Therapeutic Alliance Scales (VTAS; Hartley & Strupp, 1983) are comprised of 44 observer-rated items and was adapted from

the VPPS to be developmentally appropriate for rating alliance for youth therapy. Alliance was defined as the collaborative bond between therapist and client (Krupnick et al., 1996). Items are rated on a six-point scale ranging from 0 = *not at all* to 5 = *a great deal*.

The VTAS has demonstrated acceptable interrater reliability and internal consistency across several studies (i.e., Hartley & Strupp, 1983; Kamin, Garske, Sawyer, & Rawson, 1993; Krupnick et al., 1996) and shows convergent validity with other alliance measures (e.g., Working Alliance Inventory – Observer Report; Tichenor & Hill, 1989). Krupnick and colleagues (1996) examined the utility of the VTAS in predicting treatment outcome for patients receiving outpatient therapy for depression as part of the National Institute of Mental Health Treatment of Depression Collaborative Research Program. A revised version of the VTAS (VTAS-R) was generated by deleting items applicable to psychodynamic therapy and revising the coding manual to make it more applicable to different types of treatments (Krupnick, et al., 1996). Coders rated early (e.g., Session 3), middle (e.g., Session 9), and late (e.g., Session 15) alliance after watching videotaped sessions of patients who received either CBT, interpersonal therapy, imipramine plus clinical management, or placebo plus clinical management. Results revealed that early alliance scores and mean alliance scores (calculated by averaging early, middle, and late alliance) significantly predicted client-rated outcome (Krupnick et al, 1996).

Summary of alliance measures. In conclusion, the AOR, TPOCS-A, and VTAS have demonstrated strengths in the assessment of the alliance for youth therapy. These include promising psychometric properties of both the VPPS and TPOCS-A. Initial examination of the AOR also yields promising results. In addition, the TPOCS-A and VTAS have demonstrated predictive validity in studies examining the alliance-outcome association.

However, these measures share conceptual overlap with involvement measures such as the CIRS and SCCS, as they are adapted from the same adult-focused therapy process scale (i.e., VPPS). In fact, wording of items on these alliance measures overlap with dimensions of involvement (e.g., level of self-disclosure, behavioral involvement in therapeutic tasks, intensely negative/positive client responses). This wording makes it unclear if client behaviors are rated with a focus on the child's relationship with the therapist or the child's participation in therapeutic tasks, and creates difficulties when teasing apart the distinct effects of alliance and involvement in process research.

Chapter Four

The Present Study

Measurement of client involvement, a therapy process component that is critical to client outcomes, is inconsistent. Despite the importance of this variable across different theoretical orientations, no pantheoretical definition of client involvement exists. This definitional issue is reflected in the limited agreement in how involvement is operationalized across measures. The inconsistencies in definition and measurement create a problem when attempting to synthesize and interpret the empirical findings surrounding involvement in youth psychotherapy (Martinez et al., 2014). The current study sought to ameliorate this problem by (a) providing a definition that captures important aspects of client involvement across different treatments and orientations based on a literature review, and (b) developing a measure that will adequately assess in-session involvement based on this pantheoretical definition.

Based on these goals, a generalizability framework seemed most appropriate when examining the development and initial examination of the psychometric properties of the TPOCS-I. Within this framework, measured scores are evaluated based on their usefulness, or the extent to which the scores accurately generalize to a wider set of behaviors (Shavelson, Webb, & Rowley, 1989). In the current study, the extent to which involvement varies across different dimensions, or facets (Devellis, 2003) was examined and this approach provided information about whether these facets (e.g., client, treatment group) limit the generalizability of the TPOCS-I. In the following sections, the sample and setting in which the current study took place are described.

Choosing the Sample and Setting

Careful consideration is necessary when choosing a treatment setting and sample for the examination of in-session involvement, such as whether the sample will provide opportunities to view high and low levels of in-session involvement. Youth in-session involvement was examined with a participant sample drawn from a two larger parent studies that included youth with principal anxiety diagnoses and their therapists. The sample consists of a randomized efficacy trial and a randomized effectiveness trial and therapy took place across outpatient office, school-based clinic, and academic clinical settings. The selected dataset provides ample opportunity to observe a wide range of youth involvement activities within different treatment settings, while youth are participating in different types of psychotherapy for anxiety.

The current sample included three different conditions (or treatment approaches): (a) standard EBT with manualized instruction and prescribed order of sessions; (b) modular treatment with manualized instructions for individual treatment modules and decision-points for choosing module sequence; and (c) usual care (i.e., UC). Both modular and standardized treatments included a variety of interventions including emotion education, cognitive restructuring, and behavioral rehearsal of new skills. This made an appropriate sample for examining involvement because it afforded the opportunity to test a pantheoretical definition of involvement (i.e., behavioral, affective, and cognitive components) across three different treatment groups (Morris et al., 2014). This sample also provided the opportunity to examine variance in involvement across treatments. There is some evidence suggesting that involvement scores will not differ across treatment groups within youth suffering from symptoms of depression (Karver et al., 2008), but this has not been examined with anxious samples of youth. See the Methods section for a full description of these trials.

The current sample consists of youth with a principal anxiety diagnosis. Because anxiety disorders are highly prevalent in youth (i.e., with one in four adolescents affected in the United States; Merikangas et al., 2010), exploring the psychometric properties of in-session involvement with an anxiety-focused sample was a sound first step. Specifically, providing psychometric validation with the use of a highly prevalent disorder increases instrument utility.

In addition, youth participating in anxiety-focused psychotherapy are often asked to complete a variety of tasks within session including approaching feared stimuli/situations, identifying and modifying maladaptive thoughts, identifying affective arousal, learning skills to cope with distress, and applying these skills during situations that lead to distress. These in-session tasks require youth to participate behaviorally, affectively, and cognitively. This provided the opportunity to assess a variety of involvement behaviors and was important when the reliability of the proposed measure was assessed.

This study aimed to use sessions-by-session data to assess variance across time, treatment, settings, therapist, and client. The ability to assess youth in-session involvement variance over time is a feature needed in involvement instruments as there is some evidence to suggest that youth involvement varies across the course of treatment (Chu & Kendall, 2004; Hudson et al., 2014). In other words, measurement of involvement at one time point may provide an inaccurate picture of a child's overall involvement in therapy and lead to imprecise predictions of other variables of interest, such as outcome. Therefore, it is important that measurement take place over multiple sessions to provide an accurate picture of youth in-session involvement within therapy.

As previously noted, youth selected for the current study included those presenting with principal anxiety diagnoses. All youth meeting these criteria were included unless (a) they had

fewer than two sessions available for observational coding or (b) had more than one therapist during the course of treatment. These exclusion criteria are important because they adhere to recommendations that at least three time points be analyzed when looking at trends over time (Karver et al., 2008) and remove the potential for added variance due to having more than one therapist.

Chapter Five

Method

The current study included data and a subsample of participants drawn from two randomized clinical trials: (a) the Child System and Treatment Enhancement Projects (STEPS; Weisz et al, 2012) and (b) a sample of youth who received manualized CBT (Kendall et al., 2008).

Child STEPS

Child STEPS was a two-year randomized clinical trial that examined the effectiveness of three treatments for youth with principal anxiety, depression, and/or disruptive behavior problems (Weisz et al., 2012). Youth (N = 174) ranged in age from 7-13 and were randomly assigned to one of three treatment conditions: (a) standard treatment involving the use of one of three evidence-based treatments – Coping Cat (Kendall, 1994; Kendall, Kane, Howard, Siqueland, 1990) for anxiety, Primary and Secondary Control Enhancement Training (PASCET, Weisz et al., 2005) for depression, and Defiant Child (Barkley, 1997) for disruptive behavior; (b) modular treatment, Modular Approach to Therapy for Children (MATCH; Chorpita & Weisz, 2005), involving the use one modular treatment protocol that contained clinical procedures similar to the standard treatments, but allowed for flexibility in the delivery sequence of treatment modules; and (c) usual care. Therapists (N = 84) from 10 different outpatient community clinics and school based settings in Hawaii and Massachusetts provided treatment. Therapists saw an average of 2.07 (SD = 1.31) cases. In total, 59 youth and 29 therapists participated in the standard condition, 62 youth and 28 therapists participated in the modular condition, and 53 youth and 27 therapists participated in the usual care condition.

Youth were included in Child STEPS if they were diagnosed with principal *DSM-IV* anxiety, depressive, or disruptive behavior diagnoses as indicated through: (1) the Children's Interview for Psychiatric Syndromes (i.e., ChIPS), a structured diagnostic interview (Weller, Weller, Rooney, & Fristad, 1999a; 1999b); or (2) clinically elevated problem levels as indicated by T scores ≥ 65 on Externalizing, Internalizing, and Total Problems subscales of the Child Behavior Checklist (CBCL) and Youth Self-Report (YSR; Achenbach & Rescorla, 2001). Exclusion criteria consisted of mental retardation, pervasive developmental disorder, psychotic symptoms, principal bipolar disorder, or principal inattention or hyperactivity. Principal disorder and clinical problem for each case was identified using the CBCL, YSR, diagnoses, and the parent- and youth-identified top problems (obtained using standardized Top Problems measure via telephone interview; Weisz et al., 2012). The present study only focused on youth who had a principal anxiety disorder (N = 55). To be included in the present study, youth participants were required to have (a) a principal anxiety disorder diagnosis, (b) a minimum of two audible sessions, and (c) received treatment from a single therapist.

Participants.

***Youth participants*.** The current study sample consisted of 55 youth participants (ages 7-13 years; *M* = 9.89, SD = 1.71) assigned to receive any one of the three treatments conditions. Youth participants consisted of 58.8% male and were an ethnically/racially diverse sample, including 51.5% Caucasian, 29.4% Multiracial, 5.9% African American, 2.6% Latino, 4.4% Asian American/Pacific Islander, and 2.9% other. Annual family income was less than $60,000 for 57.3% of the families. See Table 1 for youth and family descriptive information.

***Therapist participants*.** A total of 35 therapists were included in this study. The majority of therapists were female (80.8%) and over one-third were Caucasian (44.4 %), while 22.6 %

identified as Asian American/Pacific Islander, 5 % African American, 8.1% Multiracial, and 1.6% Latino. Therapist's age ranged from 25 to 63 (M = 40.27, SD = 10.21). Therapists ranged in years of clinical experience, 0 to 30, with an average of 5.41 years (SD = 5.35). See Table 1 for detailed therapist descriptive information.

Treatments. Child STEPs compared the effectiveness of three treatments, standard manual treatment condition (EBT with manualized instruction and prescribed order of sessions), modular treatment, (MATCH, a modular treatment with manualized treatment modules and decision points for choosing model sequence; Chorpita & Weisz, 2005), and a UC condition. Therapists were randomly assigned to condition using blocked randomization stratified by the educational level of the therapist (Fayers, Jordhoy, & Kaasa, 2002).

***Standard manual treatment condition*.** Therapists randomly assigned to the standard manual treatment condition were trained to use Coping Cat (CC; Kendall, 1994; Kendall et al., 1990), a 16 to 20 session individual cognitive behavioral treatment for anxiety. A primary focus of CC is skill building aimed at the identification and remediation of unrealistic fearful thoughts, relaxation, and graduate exposure to feared object. Opportunities to practice skills are provided in session via role-play and real-life exposures and out of session via homework assignments. CC included manualized instructions for interventions and a prescribed order of treatment sessions.

***Modular treatment*.** Therapists randomly assigned to modular treatment were trained to use the Modular Approach to Treatment of Children (MATCH; Chorpita & Weisz, 2005). MATCH is a treatment containing 31 treatment modules corresponding to the treatment procedures used in CC and two other manualized treatments (i.e., Defiant Child and Primary and Secondary Control Enhancement Training, or PASCET). Therapists assigned to modular

treatment used the MATCH algorithm to choose the sequencing of module presentation and the adaptation of this sequencing should interference occur (e.g., a comorbid condition or stressor obstructs the use of the default sequencing). For example, MATCH therapists might choose to use procedures earlier than indicated, omit procedures that were not well-suited for their current youth needs, or use procedures for multiple problem areas within in the same session. MATCH focuses on building skills to help youth manage symptoms and enhance functioning, and is used to address symptoms of anxiety, depression, and/or disruptive behavior.

UC. Therapists assigned to UC practiced treatment procedures that are regularly used in this setting. Treatment continued until termination.

Therapist training. Therapists assigned to standard manual and modular treatment conditions were trained by experts in each treatment protocol. Training included two full days focused on treatment for each problem area, for a total of six days. Afterwards, both standard and modular treatment therapists received individual weekly consultation from postdoctoral project consultants. Consultants were also trained by experts in respective treatment protocols and participated in weekly discussions with experts including review of measurement feedback on client progress and practice history. Therapists assigned to UC received the typical supervision provided in their setting. Project personnel did not intervene in UC therapist procedures with the exception of routine retrieval of audiotapes of UC sessions. Treatment integrity was assessed using a checklist via observational coding of session content and revealed adherence to condition in all three conditions. In the standard condition, 93.0% of session content fit treatment elements described in the CC manual. In the modular condition, 83.0% of content fit treatment elements described in the MATCH manual. Only 92.0% of content in UC was independent of treatment elements of MATCH or CC (Weisz et al., 2012).

Diagnostic and symptom measures. The following measures were applied in STEPs to assess diagnosis and treatment outcome.

Children's Interview for Psychiatric Syndromes – Child and Parent Versions (ChIPS; Weller, Weller, Rooney, & Fristad, 1999a; 1999b). The ChIPS is a structured diagnostic interview with well-documented reliability and validity in outpatient and inpatient samples (Fristad, Glickman, Verducci, Teare, & Weller, 1998). Both youth and parent versions were administered, and diagnoses were generated by integrating youth and parent reports. Diagnoses generated by both informants were accepted as were internalizing diagnoses generated by youth informants and externalizing diagnoses generated by parent-report.

Child Behavior Checklist for Ages 6 – 18 (CBCL; Achenbach & Rescorla, 2001). The CBCL is a 113-item caregiver-report measure of youth emotional and behavioral symptoms. Each item is rated from 0 (not true) to 2 (very true or often true). This measure includes eight narrow-band subscales and three broad-band subscales. The three broadband subscales (e.g., Total, Externalizing, and Internalizing Problems) were used as indicators of clinical impairment. Respondents who obtain T scores at or above 65 exhibit clinically elevated symptoms and may be in need of treatment. The validity and reliability of this instrument has been well documented (Achenbach & Rescorla, 2001).

Youth Self-Report for Ages 11-18 (YSR; Achenbach & Rescorla, 2001). The YSR is a youth self-report form that corresponds to the CBCL and is designed to assess emotional and behavior problems in youth aged 11 to 18. The validity and reliability of this instrument has been established in multiple populations (Achenbach & Rescorla, 2001), and the broadband subscales have proven to be reliable and valid in youth as young as 7 years old (Ebesutani, Berstein,

Martinez, Chorpita, & Weisz, 2011). The Total, Externalizing, and Internalizing Problems broadband subscales were used as indicators of clinical impairment.

Kendall Coping Cat Study

The Kendall Coping Cat study (Kendall et al., 2008) was a randomized controlled trial that compared the efficacy of individual cognitive behavioral therapy (ICBT), family cognitive behavioral therapy (FCBT), and a family-based education/support/attention active control (FESA) for the treatment of youth anxiety. Youth (N = 161) ranged in age from 7-14 and were randomly assigned to one of three treatment conditions— ICBT using the Coping Cat therapist manual and workbook (Kendall & Hedtke, 2006a, 2006b); family cognitive-behavioral therapy using a family CBT manual for anxious youth (Howard, Chu, Krain, Marrs-Garcia, & Kendall, 2000) and Coping Cat workbook; and family education/support/attention using a manual for family education, support, and attention for anxious children (Krain, Hudson, Choudhury, & Kendall, 2000). Exclusion criteria consisted of psychotic symptoms, mental retardation, a disabling medical condition, the child's participating in concurrent treatment, or the child's taking antianxiety or antidepressant medications. At least one parent was required to be English speaking. Principal disorder for each case was identified using the Anxiety Disorders Interview Schedule for children (ADIS – C/P; Silverman & Albano, 1996). Therapists included master's level therapists with 2-3 years of experience at an academic training clinic specializing in anxiety disorders and doctoral-level psychologists. In total, 55 youth and participated in the ICBT condition, 56 youth participated in the FCBT condition, and 50 youth and participated in the FESA condition. Twenty therapists participated in this study, were trained in all conditions, and administered treatment.

Youth in the current study included only those youth who participated in the ICBT condition (N = 51). To be included in the present study, youth participants were required to have (a) a session from the skill building (sessions 2-8) and exposure phase (sessions 9-16) of treatment, and (b) received treatment from a single therapist.

Participants.

Youth participants. The current study sample consisted of 51 youth participants (ages 7-14 years; M = 10.35, SD = 1.89) in the ICBT condition. Youth participants consisted of 60.8% male youth and were 86.3% Caucasian, 9.8% African American, 2% Latino, and 2% other. Annual family income was less than $60,000 for 35.5% of the families. See Table 1 for youth and family descriptive information for the ICBT subsample used for this study.

Therapist participants. A total of 16 therapists were included in this study. The majority of therapists were female (88.2%) and over half were Caucasian (68.6%), while 9.8% identified as Asian American/Pacific Islander, and 17.6% Latino. See Table 1 for detailed therapist descriptive information.

Individual Cognitive Behavioral Treatment. The Coping Cat program was delivered to address youth anxiety problems (Kendall & Hedtke, 2006a; Kendall & Hedtke, 2006b). A primary focus included skill building aimed at the identification and remediation of unrealistic fearful thoughts, relaxation, and graduate exposure to feared object. Youth were taught to manage anxiety using the FEAR acronym: (a) recognizing anxious feelings and associated somatic reactions (i.e., *Feeling frightened?*), (b) identifying anxious thoughts (i.e., *Expecting bad things to happen?*), (c) developing a coping plan that included modifying anxious thoughts and engaging in coping behavior (i.e., *Actions and attitudes that can help*), (d) evaluating efforts and self-reward (i.e., *Results and rewards*). Opportunities to practice skills were provided in session

via role-play and real-life exposures and out of session via homework assignments. Therapists provided youth and parents with education about youth anxiety and coached parents on how to respond to their youth's anxious behavior.

Therapist training. All therapists studied written therapy manual materials and participated in training that generally consisted of two 3-hour workshops. Workshops included didactic presentation, role plays, trainee demonstration, videotape playback, and discussing. Therapists participated in weekly two-hour group supervisions following training. Treatment integrity was rated using a checklist that contained the content/strategies called for in sessions by the manual. Experienced therapists rated 14.0% of sessions and indicated whether the appropriate content was covered. Results showed that 91.0% of content was appropriate in the ICBT condition.

Diagnostic and symptom measures. The following measures were utilized to assess diagnostic and symptom domains.

Anxiety Disorders Interview Schedule for Children *(ADIS-C/P; Silverman & Albano, 1996).* The ADIS-C/P is a semi-structured diagnostic interview with well-documented reliability and convergent validity. Both youth and parent versions were administered. Anxiety diagnoses were assigned if either the youth or parent reported the diagnosis and the clinical assigned a clinical severity rating of four or greater (with 0 = not at all, 4 = some, and 8 = very, very much).

Child Behavior Checklist the CBCL, described previously, was utilized as a parent report form for assessing youth mental health symptoms.

Development of the Therapy Process Observational Coding System for Child Psychotherapy– In-Session Involvement Scale (TPOCS-I)

The TPOCS-I is a 14-item comprehensive coding system designed to measure youth involvement in therapy. The TPOCS-I was developed via a four-step process.

Step 1: Subscale focus. The first step involved reviewing the literature to identify the involvement dimensions relevant to youth therapy to create subscales for the involvement measure. The three dimensions proposed by Morris and colleagues (2014), behavioral, affective, and cognitive involvement, were selected for this measure as they had good conceptual support. See Table 1 for items from each subscale. Behavioral involvement includes learning and practicing new skills such as relaxation, assertiveness, social skills, and anger management techniques. Affective involvement includes engaging in activities where emotions are experienced and processed, such as labeling and accepting emotions, mindfulness exercises, body scanning, and exposure. Cognitive involvement includes exploring and commenting on thought processes such as learning about distorted styles of thinking, identifying distorted thoughts, and altering distorted thoughts into more adaptive ones. Positive involvement includes actively participating in the therapy process across behavioral, affective, and cognitive domains. Negative involvement includes demonstration of passivity or resistance/refusal to participating in the therapy process (Chu & Kendall, 2004).

Step 2: Item development. The second step involved generating items to assess the behavioral, affective, cognitive, positive, and negative dimensions. Items were generated by evaluating the conceptual and theoretical literature on measurement of client involvement (see review in the section "Involvement focused measures"). In addition, measures of youth therapy process were reviewed because of their inclusion of items related to in-session involvement and importance in development of existing youth involvement measures (see review in the section "Measures of therapy process"). Appendix 1 also contains a table with summary information

about these measures. Items related to client involvement (n = 35) were identified. This pool of items was reduced by eliminating or combining redundant items (9 items); eliminating items that assessed motivation, child-therapist relationship, and other variables related to, but not a part of in-session involvement (10 items); and eliminating items that appeared to developmentally inappropriate for youth therapy (4 items; e.g., concern was how to deal more effectively with self and others). For example, the items "Does the child initiate discussion or introduce new topics of conversation?" and "Took initiative in bringing up the subjects that were talked about." were combined into "Does the client initiate discussions or introduce new topics of conversation related to therapeutic activities?" The resulting item pool was reviewed by two psychology faculty and two graduate students who held clinical expertise in youth therapy at Virginia Commonwealth University. On the basis of this feedback, items were edited for clarity. The final version of the involvement measures consisted of a total of 14 items. Items were either indicative of positive (9 items) or negative (5 items) involvement. Subscales contained at least one positive and one negative item and included the behavioral (6 items), affective (3 items), and cognitive (5 items) subscales. A coding manual was developed to assist coders in accurate and reliable coding of each item with descriptions, exemplars, and instructions for differentiation between items. See Table 2 for a list of involvement measure items, and their sources.

Step 3: Scoring strategy. A scoring strategy was determined after the items were finalized. The CIRS 6-point Likert-type scale was used for ratings with the following anchors: 0 = not at all, 2 to 3 = somewhat, and 5 = a great deal. The scoring strategy involved (a) ratings on a 6-point Likert-type scale, (b) ratings of the entire treatment sessions, and (c) ratings of youth. An even number of response options was chosen to optimize variability and decrease the likelihood that respondents would over-rely on a middle or neutral response (DeVellis, 2003).

Step 4: Pilot coding. The measure was used to pilot code approximately 40 therapy sessions. Interrater reliability of each item was assessed, and those items that demonstrated low reliability (intraclass correlation [ICC] below .40; Cicchetti, 1994) were refined. ICCs were based on two-way random effects model for the average of two coders. During the piloting phase, coders provided feedback on item content, definitions, and exemplars. Feedback was used to refine the items and manual.

Measures Used for Validity Analyses

The following observational measures were used in to assess construct validity of the proposed measure of involvement.

Child Involvement Rating Scale (CIRS; Chu & Kendall, 2004). The CIRS consists of 6 items related to child involvement in session. CIRS data were collected for the Kendall study only. Coders observed entire sessions and rated each item on a six-point scale from 0 = *not at all* to 5 = *a great deal*. The CIRS has demonstrated adequate interrater reliability ranging from .76 to .90 and internal consistency of .73 (Chu & Kendall, 2004, 2009). Interrater reliability within the current study fell within the good to excellent ranges (ICC = 0.68 – 0.81) and internal consistency is alpha = 0.85..

Cognitive Behavioral Treatment for Anxiety in Youth Competence Scale (CBAY-C; McLeod et al., 2016). Therapist competence was measured using an observational measure designed to capture therapist limited-domain competence (i.e., competence in the delivery of core interventions and delivery found in a specific psychosocial treatment program) and delivery of core practice elements in ICBT. This 25-item measure is made up of five standard (i.e., interventions common to CBT across each session), 12 model (i.e., theory driven interventions specific to CBT for youth anxiety and expected to be focus in one or more sessions such as

exposure), six delivery (i.e., how specific interventions were delivered such as via role-play), and 2 global items that focus on overall skillfulness and responsiveness in therapist intervention delivery. Coders rated each item on the extent to which a therapist intervention delivery quality (i.e., skillfulness) and appropriateness of delivery for given client and situation (i.e., responsiveness) using a 7-point Likert-type scale with the anchors 1 = *very poor*, 3= *acceptable*, 5 = *good*, and 7 = *excellent.* The CBAY-C Scale, Subscale, and item scores have demonstrated evidence of construct validity (McLeod et al., 2016). The mean inter-rater reliability for the CBAY-C items in the current study was ICC (2,2) = 0.69 (*SD* = 0.11).

Therapy Process Observational Coding System for Child Psychotherapy-Alliance Scale (TPOCS-A; McLeod & Weisz, 2005). The TPOCS-A is a nine-item scale that assesses the quality of the client-therapist alliance in youth therapy. This measure consists of six items that assess affective elements of the client–therapist relationship, and three items that assess client participation in therapeutic activities. Coders observed entire sessions and rated each item on a six-point scale ranging from 0 = *not at all* to 5 = *a great deal*. The TPOCS-A has demonstrated item interrater reliability ranging from .48 to .80 (*M* ICC = .67), internal consistency ranging from .91 to .95 (*M* α = .92), and convergent validity with self-report alliance measures ranging from .48 to .53 (Fjermestad et al., 2012; Liber et al., 2010). Interrater reliability for the TPOCS-A in the current study was ICC = 0.86 and internal consistency was alpha = 0.89.

Therapy Process Observational Coding System for Child Psychotherapy-Revised Strategies Scale (TPOCS-RS; McLeod, Smith, Southam-Gerow, Weisz, & Kendall 2015). The TPOCS-RS assesses the extensiveness of interventions delivered in therapy. It consists of 42 items that comprise five subscales: Cognitive (4 items; e.g., Cognitive Distortion), Behavioral (9

items; e.g., Operant Interventions), Psychodynamic (5 items; e.g., Interpretation), Family (7 items, e.g., Parenting Intervention), and Client-Centered (4 items; e.g., Positive Regard). In addition, there are 13 items (e.g., Homework, Play Therapy) that represent therapeutic interventions that play a meaningful role in therapy but are not associated with a specific subscale. Coders observed entire sessions and rated the extent to which a therapist engages in each intervention during that session using a 7-point rating scale with the following anchors: 1 = *not at all*, 3 = *somewhat*, 5 = *considerably,* and 7 = *extensively*. The TPOCS-RS item scores have demonstrated adequate interrater reliability (ICC > .60) and there is support for construct validity (McLeod et al., 2015; McLeod & Weisz, 2010; Southam-Gerow et al., 2010). For the current study, interrater reliability was calculated for the following subscales: Cognitive (ICC= 0.89), Behavioral (ICC = 0.89), and Client Centered (ICC = 0.61). Interrater reliability was also calculated for one item, Encourages Affect (ICC = 0.57).

Vanderbilt Therapeutic Alliance Scale - Revised (VTAS-R Short Form, Shelef & Diamond, 2008). The VTAS-R Short Form is a five-item observer-rated scale based on the Vanderbilt Therapeutic Alliance Scale (VTAS-R; Diamond et al., 1999). The VTAS-R Short Form measures the strength of the therapeutic alliance in individual therapy for youth. Items were rated on a six-point scale ranging from 0 = *not at all* to 5 = *a great deal.* Initial psychometric data reveals good to excellent interrater reliability with ICCs ranging from .72 to .87 and high internal consistency with coefficient alphas at .90 and .91 for adolescent and parent observations, respectively (Shelef & Diamond, 2008). For the current study, interrater reliability was ICC = 0.82 and internal consistency was alpha = 0.88.

Observational Scoring and Session Sampling Procedures

To ensure that coders were properly trained and to minimize rater drift, the following procedures were used for observational coding.

Sampling of treatment sessions. All available recordings were used for observational coding for the Child STEPS sample unless they met the following exclusion criteria. Sessions were not rated if the session (a) was less than 15 minutes in length, (b) had less than 15 minutes audible, (c) contained less than 75.0% of English dialogue, or (d) audio/video file(s) were missing or damaged. The first and last sessions for each client were also excluded, as the processes conducted in these sessions might impact client involvement (Windholz, Weiss, & Horowitz, 1985). In all, 797 Child STEPS sessions were coded.

For the Kendall sample, 98 sessions were coded. Sessions were selected from the ICBT arm of the Kendall sample. Sessions from each case were selected from the first half and the last half of treatment. The first half of treatment includes building skills such as relaxation and problem solving related to anxiety management. The second half of treatment emphasizes exposure. Sessions were not rated if the session was (a) was less than 15 minutes in length, (b) had less than 15 minutes audible, (c) contained less than 75.0% of English dialogue, or (d) audio/video file(s) were missing or damaged. The first and last sessions were also excluded.

Coders. All coders were graduate students in a doctoral clinical psychology program. The Child STEPS coders consisted of five female graduate students. Mean coder age was 27.6 years (SD = 1.34), and 60.0% of coders self-identified as Caucasian, 10.0% as Asian, and 10.0% as multiracial. Three coders rated the alliance (TPOCS-A) and therapist competence (CBAY-C). Coders (*M* age = 27.33; *SD* = 1.53) included three female graduate students; 67.0% of coders self-identified as Caucasian and 33.0% identified as Asian. Two (one male, one female) graduate students (*M* age = 30.50; *SD* = 3.54), assessed client involvement (CIRS) in the Kendall sample.

All coders had previous coding experience. In addition, each coder had at least one year of clinical experience, and had been trained in treatment with children and adolescents.

Coder training. Coders were trained over a two-month period to reach adequate pre-study reliability at the item level (ICC > .59; Cicchetti, 1994). Training consisted of reading the scoring manual, reviewing specific session segments, and practicing scoring sessions. Coders completed 40 certification sessions to demonstrate adequate reliability, and afterwards coders were allowed to rate sessions independently. Once scoring began, sessions were randomly assigned to coders, and regular reliability assessments were performed. The results of these assessments were discussed in regular weekly meetings to prevent coder drift. Frequency of these meetings decreased over time as coders continued to demonstrate adequate reliability. Coders were naive to therapy outcome, site, and condition.

Scoring of therapy sessions. Coders scored entire treatment sessions, which ranged in length from 15 minutes to 112 minutes (M = 41.20; SD = 47.90). Each session was double-coded for reliability. For process analysis, mean scores of each item were used instead of using scores produced by one coder, as this has been shown to reduce measurement error by removing differences among coders (Lambert & Hill, 1994; McLeod et al., 2013).

Data Analytic Strategy

The purpose of the current study was to develop a measure using a pantheoretical definition that focused on behavioral, affective, cognitive, positive, and negative domains (Morris et al., 2014) of in-session client involvement and investigate the initial score reliability and validity of the TPOCS-I. To achieve this goal, data analysis focused on evaluating: (a) TPOCS-I item performance, (b) TPOCS-I scale and subscale performance, and (c) the extent to which the TPOCS-I behavioral, affective, cognitive, positive, and negative subscales performed

as anticipated. To achieve study goals, two samples were used (i.e., Child STEPS, Kendall). Analyses were conducted separately for the two samples to assess if the TPOCS-I performed consistently across the samples. The TPOCS-I is intended to be a pantheoretical measure of in-session child involvement that is applicable across diagnoses and treatment approach. It was therefore important to examine the pattern of findings in two different samples to assess how well this measure generalized to youth presenting with diagnostic diversity (i.e., anxiety and comorbid conditions) and receiving varying therapeutic approaches.

Preliminary analyses. Prior to assessing the score reliability and validity of the TPOCS-I items, subscales, and scales a series of preliminary analyses were conducted. First, youth and therapist demographic information was compared to the parent samples. These analyses were conducted to determine the extent to which conclusions from this sample could be generalized to the parent studies. Second, the Kendall and Child STEPS samples were compared to see if there were differences in youth, family, and therapist variables.

Evaluation of TPOCS – I scale and subscale scores

A series of steps were conducted to transition the TPOCS-I items to subscales. This work proceeded using a pantheoretical definition of involvement. Morris and colleagues (2014) proposed that involvement consists of behavioral, affective, and cognitive components whereas Chu and Kendall (2004) focused on positive and negative components. Items were developed for each subscale (see Table 2 for a list of items hypothesized to fit each subscale). Analyses proceeded through four steps to determine if the items performed as designed and mapped onto each subscale. Findings from these analyses were used to inform the development of the subscales for the validity analyses.

Step 1: Inter-rater reliability. Inter-rater reliability of each item was assessed using intraclass correlation coefficients (ICCs; Shrout & Fleiss, 1979). The reliability coefficients represent the model ICC (2,2) based on a two-way random effects model. The ICC provides an estimate of the ratio of the true score variance to total variance. These correlations therefore provide a reliability estimate of the mean scores of all coders. This ICC model also allows for generalizability of the results to other samples. Following the recommendations of Cicchetti (1994), ICC values below .40 reflect "poor" agreement, ICCs from .40 to .59 reflect "fair" agreement, ICCs from .60 to .74 reflect "good" agreement, and ICCs .75 or higher reflect "excellent" agreement.

Descriptive statistics of each TPOCS-I item were examined including frequency, range, mean, and standard deviations for each item. TPOCS-I item scores were created by averaging the scores produced by each coder. The descriptive data were interpreted using DeVellis' (2003) guidelines that recommend for individual items to capture different levels of a construct the items should demonstrate a mean close to the center of the mid-point on the scale, use the full range of the scale, and minimal skew and kurtosis. Items with ICC values of .40 or lower were analyzed at this point in an attempt to identify the reason for the low inter-rater reliability (e.g., restricted range).

Step 2: Internal consistency. The internal consistency of each subscale was evaluated to assess the reliability of the subscales as well as how internal consistency was impacted by the removal of specific items. Items that were negatively valenced were reverse coded: (a) *Withdrawn*, (b) *Inhibited/Avoidant*, (c) *Distracted*, (d) *Oppositional*, and (e) *Passive*. Internal consistency of the Total and subscale scores (i.e., Behavioral, Affective, Cognitive, Positive,

Negative) was examined using Cronbach's Alpha, with a target α between .80 and .90, as recommended by Clark and Watson (2003).

Step 3: Construct validity of the TPOCS-I item scores. I next evaluated whether the TPOCS-I item scores demonstrated evidence of construct validity. These analyses focused on the magnitude of the correlation among the TPOCS-I items. Correlations were judged based on Rosenthal and Rosnow's (1984) standards where *r* of 0.10 to 0.23 indicates a "small" effect, 0.24 to 0.36 indicates a "medium" effect, and an *r* greater than 0.36 indicates a "large" effect. It was expected that inter-item correlations would be consistent with subscale configuration such that items within each subscale would share more conceptual overlap (i.e., "medium" to "large") and would demonstrate "small" to "medium" overlap with items across subscales (DeVellis, 2003).

Step 4: Subscale creation. Results from the analyses focused on the items, subscales, and total scale were used to inform the creation of the subscales for the following analyses focused on the score validity of the TPOCS-I scale and subscales. Scale and subscale scores were created by averaging the items that comprised each scale or subscale.

TPOCS-I scale and subscale validity studies

Construct validity of the TPOCS-I scale and subscale scores was assessed and this information was used to evaluate whether the proposed subscales performed as expected. The goal of these analyses was to determine whether the proposed scale and subscale scores related to scores on other measures in predicted ways.

Construct validity. Construct validity is the degree to which a measure examines what it claims to assess (DeVellis, 2003). One way to assess construct validity is to assess the extent to which two variables that are proposed to share conceptual variance overlap statistically. This is an examination of convergent validity. Another way to assess construct validity is to assess

whether variables that are supposed to differ conceptually are in fact unrelated. This is an examination of discriminant validity. To examine construct validity for the TPOCS-I, Pearson correlations were used to examine the relations between the TPOCS-I scale and subscales and other measures that have been theoretically and empirically linked to in-session involvement in the past.

First, relations between the TPOCS-I Total and Subscales (Affective, Behavioral, Cognitive, Positive, Negative) scores and a measure of child involvement (CIRS) were evaluated using the Kendall sample. Scores on the CIRS and TPOCS-I were produced by a separate team of coders. It was expected that these measures would have positive correlations within the "large" range as they are both designed to measure in-session involvement.

Second, correlations between the TPOCS-I Total and Subscale (Affective, Behavioral, Cognitive) scores and measures of child alliance (TPOCS-A and VTAS-R) were evaluated using the Kendall and Child STEPS samples. Conceptually, client involvement and the alliance are considered to be distinct aspects of the therapy process (Fjermestaad et al., 2016); however, past efforts to assess client involvement and the alliance with observational measures have raised questions about the ability of coders to distinguish between these two constructs (McLeod et al., 2012). So, evaluating the degree of overlap between the TPOCS-I Total and Subscale scores and the two alliance measures (i.e., TPOCS-A and VTAS-R) represented an important test of the discriminant validity of the TPOCS-I Total and Subscales.

Third, the association between the TPOCS-I Total and Subscale (Behavioral, Affective, Cognitive, Positive, Negative) scores and the quality of intervention delivery (i.e., therapist competence) was examined. Given previous findings of correlations between therapist competence, defined as flexible and competent intervention delivery, and client involvement in

the “medium” range (Hudson et al., 2014), it was expected that the correlations between therapist competence and the TPOCS-I scale and subscale scores would also fall in the “medium” range.

Last, associations between the TPOCS-I Behavioral, Affective, Cognitive, Positive, and Negative subscales and theory-based domains of therapist interventions (i.e., TPOCS-RS Cognitive and Behavioral subscales and Encouraging Affect item) were examined. It was expected that “matched” subscales (e.g., TPOCS-RS Behavioral subscale and TPOCS-I Behavioral subscale) would yield correlations falling in “medium” range given that therapist delivery of specific interventions is anticipated to provide more opportunities and prompting for client in-session involvement behaviors related to these interventions. Therefore, the following correlations were expected to fall in the “moderate” range: (a) TPOCS-RS Behavioral subscale and TPOCS-I Behavioral subscale; (b) TPOCS-RS Cognitive subscale and the TPOCS-I Cognitive subscale; (c) TPOCS-RS Encouraging Affect item and TPOCS-I Affective subscale. It was anticipated that “non-matched” subscales (e.g., TPOCS-RS Behavioral and TPOCS-I Cognitive) would yield “small” correlations. In addition, it was anticipated that correlations between the TPOCS-I Total score and TPOCS-RS Cognitive and Behavioral subscales and Encourages Affect item would fall in the “small” range.

Variance components analysis. Variance components analysis examines possible sources of variance attributable to coders, therapists, clients, time, setting, and residual error. Variance components analysis estimates how much of the variance in involvement is attributable to each of these domains (e.g., coders and therapists). Linear mixed-effects model procedures with restricted maximum likelihood estimations were be used to estimate variance components to determine the proportion of variance in the proposed measure accounted for by potential sources of variance in coders, therapists, clients, time, setting, and residual error.

Chapter Six

Results

Preliminary Analyses.

Data from the Kendall (Kendall et al., 2008) and Child STEPS (Weisz et al., 2012) samples were examined to identify if the sample from the current study was representative of the overall sample. There were no significant differences for youth age, sex, or race/ethnicity; family income; or therapist sex, race/ethnicity, age, or years of professional experience following degree for the Kendall and Child STEPS sample.

Reliability Analyses.

Step 1: Inter-rater reliability. Inter-rater reliability, ICC (2,2), for the TPOCS-I items ranged from 0.16 to 0.86 (M = 0.73, SD = 0.08) for the Kendall sample. According to Cicchetti (1994), one item fell in the "poor" range (i.e., *Inhibited*), six items fell in the "fair" range (i.e., *Initiate, Self-disclose, Ask questions, Focus, Explores feelings, Passive*), six items fell in the "good" range (i.e., *Enthusiasm*, *Cognitive*, *Behavioral*, *Participation*, *Distraction*, *Oppositional*), and one item fell in the "excellent" range (i.e., *Withdrawn*).

Inter-rater reliability, ICC(2,2), for the TPOCS-I items ranged from 0.49 to 0.78 (M = 0.82, SD = 0.08) for Child STEPS. Three items fell in the "fair" range (i.e., *Focus*, *Explores feelings*, and *Distraction*), 10 fell in the "good" range (i.e., *Initiate, Enthusiasm, Self-disclose, Ask questions, Cognitive*, *Behavioral*, *Participation*, *Inhibited*, *Oppositional*, *Passive*) and one fell in the "excellent" range (i.e., *Withdrawn*).

Next, mean item scores were produced by averaging the items scores for each coder. Descriptive information was produced for each TPOCS-I item (see Tables 3 and 4). Across both samples most items used the full, or close to, the full range of scores. Most items evidenced

acceptable skewness and kurtosis. However, three items evidenced elevated skewness (i.e., greater than 1.5) and kurtosis (i.e., greater than 1.5) across both samples: *Withdrawn*, *Distraction*, and *Oppositional*. Data from ICC's and item descriptive information were two measures of empirical performance that were used when determining whether to keep or remove items.

Step 2: Internal consistency. Cronbach's alpha was calculated as a measure of internal consistency (see Tables 7 and 8). Total 12-item scale alpha scores were 0.87 for the Kendall sample and 0.90 for the Child STEPS sample. Subscale alpha scores ranged from 0.54 to 0.70 for the Kendall sample with the Behavioral subscale at 0.65, the Affective subscale at 0.57, and the Cognitive subscale at 0.72. When *Distraction* and *Oppositional* were dropped from the Behavioral subscale, alpha was at 0.56. Subscale alpha scores ranged from 0.61 to 0.87 for the Child STEPS sample with the Behavioral subscale at 0.59, the Affective subscale at 0.67, and the Cognitive subscale at 0.82. When *Distraction* and *Oppositional* were dropped from the Behavioral subscale, alpha was at 0.61.

Step 3: Construct Validity of the TPOCS-I item scores. Inter-item correlations were calculated to examine the overlap between each TPOCS-I item (see Tables 5 and 6 for correlations from each sample) and this was another was used to provide empirical evidence when deciding whether to retain or remove items based on degree of overlap. Correlations also provided conceptual evidence to support removing or retaining items as direction of correlation revealed whether items performed as conceptually anticipated. Overall, inter-item correlations did not perform as expected as no subscale yielded correlations demonstrating the anticipated "large" effects and instead range of effects were variable. The correlations among the items on the Behavioral subscale ranged from r = -.20 to 0.66 for the Kendall sample and r = -0.15 to 0.61

for the Child STEPS sample. The *Distraction* and *Oppositional* items negative correlations across both samples with other items within the Behavioral subscale. Therefore, *Distraction* and *Oppositional* were flagged as potential items to discard as correlation evidence indicated that these items failed to converge with other items on the same subscale given the low correlations. In addition, the correlations between *Distraction* and some other items did not perform as hypothesized (e.g., negative correlations when it was expected that correlations would be positive). The correlations among items on the Affective subscale ranged from $r = 0.18$ to 0.61 for the Kendall sample and $r = 0.23$ to 0.55 for the Child STEPS subscale. The correlations among items on the Cognitive subscale ranged from $r = 0.19$ to 0.68 for the Kendall sample and $r = 0.45$ to 0.81 for the Child STEPS sample. *Initiate* and *Self-disclose* were highly correlated within the Child STEPS sample ($r = 0.81$), indicating potential redundancy. Last, four items (i.e., *Initiate, Enthusiasm, Self-disclose, Participation*) displayed high interitem correlations despite falling in different subscales.

Step 4: Subscale Creation. The next step was to use the previous data to make decisions about subscale composition. Item performance was reviewed to determine if items should be dropped. The decision to remove items was made using both statistical and conceptual evidence. *Oppositional* and *Distraction* were dropped because of poor statistical (i.e., high skewness and kurtosis) and conceptual performance (i.e., failing to converge with other items on the same subscale) across Kendall and Child STEPS samples. Some other items (i.e., *Focus; Explores Feelings*) were considered for removal because ICC's fell below the desired range for inter-rater reliability across both samples (i.e., fair range) or due to high inter-item correlations (i.e., *Initiate, Enthusiasm, Self-disclose, Participation*) because of generally high correlations with one

another. However, these items were retained because they were deemed to be a conceptually important to each respective subscale.

Following examination of the TPOCS-I items the subscales were created. Similar to item evaluation, the TPOCS-I subscale descriptive information was reviewed. Subscales included the Total 12-item scale and the proposed subscales (i.e., Behavioral, Affective, Cognitive, Positive, and Negative). Tables 7 and 8 include scale and subscale descriptive information. Data calculations were made after the removal of *Distraction* and *Oppositional*.

Correlations between the TPOCS-I subscales are presented in Table 9. Overall, the subscales demonstrated high overlap, with values in the "large" range across subscale scoring strategies (i.e., Behavioral/Affective/Cognitive and Positive/Negative) in both samples. For the Kendall sample the correlations between the Cognitive and Affective subscales were the highest, r (93) = 0.79, $p < 0.000$, followed by the correlations between the Behavioral and Affective subscales, r (93) = 0.63, $p < 0.000$, and the Behavioral and Cognitive subscales, and r (93) = 0.62, $p < 0.001$. The correlation between the Positive and Negative subscales was r (93) = 0.66, $p < 0.001$.

For the Child STEPS sample, correlations were highest between the Cognitive and Affective subscales, r (732) = 0.82, $p < 0.000$, followed by the correlation between the Affective and Behavioral subscales, r (93) = 0.72, $p < 0.001$, and the correlation between the Cognitive and Behavioral subscales, r (93) = 0.71, $p < 0.001$. The correlation between the Positive and Negative subscales was r (732) = 0.71, $p < 0.001$.

Due to the removal of the *Distraction* and *Oppositional* items based on poor item empirical performance, it was decided that the Negative subscale would not be included in the validity analyses. Without the Negative subscale, the Positive and Negative subscale

configuration was no longer deemed desirable so the Positive subscale was not included in validity analyses.

Validity Analyses

Unless otherwise noted, correlations were examined separately by participant samples (i.e., Kendall sample and Child STEPS sample).

Construct validity. Construct validity was assessed by examining if the TPOCS-I Total score and subscale scores (i.e., Behavioral, Affective, and Cognitive) performed as predicted when correlations with other related variables were examined.

Involvement and alliance. Correlations between the TPOCS-I and measures of involvement (i.e., CIRS) and alliance (i.e., TPOCS-A and VTAS) were examined (see Table 10). The correlation between the TPOCS-I Total Score and CIRS for the Kendall sample was in the large range, r (93) = 0.72, $p < 0.000$.

Correlations between the TPOCS-I Total score and the TPOCS-A were also in the "large" range for both the Kendall, r (93) = 0.62, p < 0.001, and Child STEPS, r (732) = -.73, p < 0.001, samples. The correlations between TPOCS-I Total score and the VTAS were in the "large" range for the Kendall sample, r (93) = 0.72, $p < 0.001$ and Child STEPS sample r (732) = 0.76, $p < 0.001$.

In regard to subscales, the correlations between the alliance and the TPOCS-I Behavioral, Affective, and Cognitive subscales were large. For the Kendall sample, the correlations between the TPOCS-A and the Behavioral, Affective, and Cognitive subscales were r (93) = 0.54, $p < 0.001$, r(93) = .67, $p < 0.001$, and r(93) = .51, $p < 0.001$, respectively. For the Child STEPS sample, correlations between the TPOCS-A and the Behavioral, Affective, and Cognitive subscales were r (732) = 0.62, $p < 0.001$, r (732) = 0.73, $p < 0.001$, and r (732) = 0.68, $p < 0.001$,

respectively. The correlations between the VTAS and the Behavioral, Affective, and Cognitive subscales for the Kendall sample were r (93) =0.63, $p < 0.001$, r (93) = 0.67, $p < 0.001$, and r (93) = 0.63, $p < 0.001$, respectively. For the Child STEPS sample, correlations between the VTAS and Behavioral, Affective, and Cognitive subscales were r (732) = 0.65, $p < 0.001$, r (732) = 0.71, $p < 0.001$, and r (732) = 0.72, $p < 0.001$, respectively.

Competence. Correlations between the TPOCS-I Total Score and a measure of therapist competence, the CBAY-C, were examined. Correlations between the TPOCS-I 12-item Total Score was in the "medium" range for the Kendall sample, r (93) = 0.26, $p = 0.011$, and the Child STEPS sample, r (732) = 0.28, $p < 0.001$. For the Kendall sample, correlations between the CBAY-C and the Behavioral and Affective subscales were r (93) = 0.25, $p = 0.017$, and r (93) = 0.31, $p = 0.002$, respectively. The correlation between the CBAY-C and Cognitive subscale was not significant, r (93) = 0.18, $p = 0.78$. For the Child STEPS sample, the correlations between the Behavioral, Affective, and Cognitive subscales were r (732) = 0.25, $p < 0.001$, r (732) = 0.28, $p < 0.001$, and r (732) = 0.25, $p < 0.001$, respectively.

Behavioral interventions. Correlations between the TPOCS-I Behavioral, Affective, and Cognitive subscales and the Behavioral and Cognitive subscales and the Encouraging Affect item from the TPOCS-RS were examined. Correlations between the TPOCS-RS Behavioral intervention subscale and the TPOCS-I Behavioral involvement subscale were in the "small" range for the Kendall, r (93) = 0.21, $p = 0.004$, and Child STEPS, r (732) = 0.23, $p < 0.000$, samples. Likewise, correlations between the TPOCS-RS Behavioral intervention subscale and the TPOCS-I 4-item Behavioral involvement subscale were r (93) = 0.24, p = 0.020 for the Kendall sample and r (732) = 0.22, p < 0.000. The association between the TPOCS-RS Behavioral subscale and the TPOCS-I Affective and Cognitive subscales, r (732) = 0.10, p =

0.005 and, r (732) = 0.17, $p < 0.000$, respectively, were in the "small" range. Neither correlation was significant for the Kendall sample. For the Kendall sample, the correlation between the TPOCS-I Total Score and the TPOCS-RS Behavioral subscale was not significant, r (93) – 0.03, $p = 0.754$. For the Child STEPS sample, the association was significant, but below the "small" range, r (734) = 0.09, $p = 0.012$.

Affective interventions. The association between the TPOCS-RS Encourages Affect item and the TPOCS-I Affective subscale was not significant for the Kendall sample, r (93) = 0.19, $p = 0.062$ and was in the "small" range for the Child STEPS sample, r (732) = 0.11, $p = 0.002$. Likewise, the Encourages Affect item and the Behavioral involvement subscale correlation was not significant for the Kendall sample, r (93) = 0.06, $p = 0.598$. The association was significant for the Child STEPS sample, r (732) = 0.09, $p = 0.018$, although too low to fall within the "small" range. The associations between the Encourages Affect intervention item and the Behavioral involvement 4-item subscale displayed a similar pattern; for the Kendall sample, this association was not significant, r (93) = 0.08, $p = 0.439$ and was in the "small" range for the Child STEPS sample, r (732) = 0.10, $p = 0.005$. The association between Encourages Affect and Cognitive involvement was in the "small" range for both samples with values at r (93) = 0.21, $p = 0.038$ and r (732) = 0.12, $p = 0.001$, for Kendall and Child STEPS, respectively. For the Kendall sample, the correlation between the TPOCS-I Total Score and the Encourages Affect item was not significant, r (93) = 0.19, $p = 0.069$. For the Child STEPS sample, the association was in the "small" range, r (734) = 0.12, $p = 0.001$.

Cognitive interventions. Correlations between the TPOCS-RS Cognitive subscale and the TPOCS-I Cognitive subscale were in the "large" range for the Kendall sample, r (93) = 0.39, $p < 0.000$, and the "medium" range for the Child STEPS sample, r (732) = 0.31, $p < 0.000$. The

correlations between the TPOCS-RS Cognitive subscale and the TPOCS-I Affective subscale fell in the "medium" range for the Kendall sample, r (93) = 0.28, p = 0.006, and the Child STEPS sample, r (732) = 0.28, $p < 0.000$. The correlation between the TPOCS-RS Cognitive subscale and the TPOCS-I Behavioral subscale was in the "small" range for the Child STEPS sample, r (732) = 0.19, $p < 0.000$, and was not significant for the Kendall sample, r (93) = 0.02, p = 0.888. Similarly, the correlations between the TPOCS-RS Cognitive subscale and the TPOCS-I Behavioral 4-item subscale fell in the "small" range for the Child STEPS sample, r (732) = 0.17, $p < 0.000$ and was not significant for the Kendall sample, r (93) = 0.02, p = 0.878. For the Kendall and Child STEPS samples, the correlation between the TPOCS-I Total Score and the TPOCS-RS Cognitive subscale was in the "medium" range, r (93) = 0.28, p = 0.006 and r (734) = 0.28, $p < 0.001$, respectively.

Variance components analysis. Variance components analysis on the TPOCS-I Total scale and subscale scores was conducted using SAS/STAT Software 9.4. Variance components were calculated using a mixed model with restricted maximum likelihood estimation for the following factors: (a) study group, (b) therapist, (c) youth, (d) time, and (e) coder. These factors each represent a potential source of variation in treatment delivery (Barber et al., 2004). "Study group" refers to the three different treatment groups in Child STEPS (i.e., standard manualized treatment; modular manualized treatment; and UC) and their influence on TPOCS-I scale and subscale scores. The term "therapist" represents systematic differences across therapists on the TPOCS-I scale and subscale scores. "Youth" refers to the systematic differences in TPOCS-I Total and subscale scores across each youth. "Week" represents the effect time in treatment (as measured by week) has on TPOCS-I scale and subscale scores. "Coder" reflects systematic differences in coder ratings on TPOCS-I scale and subscale scores. A separate analysis was run

for the Total 12-item scale and each subscale score. The results of the variance components analyses are presented in Table 11.

Variance in coders' ratings of the TPOCS-I Total score was accounted for primarily by residual error (0.39), followed by youth (0.33), time in treatment (0.11), and coder (0.08) for the Kendall sample. Similarly, variance in coder ratings for the Behavioral, Affective, and Cognitive subscales was primarily due to residual error (0.48, 0.55, and 0.43 respectively) followed by youth (0.31, 0.25, and 0.28 respectively). Variance attributable to coder, therapist, and time in treatment all accounted for less than 10% of the variance with the exception of the Affective subscale (coder = 0.10) and the Cognitive subscale (week = 0.22).

For the Child STEPS sample, variance in coders' ratings of the TPOCS-I Total score was accounted for primarily by youth (0.36), followed by residual error (0.35), week (0.15), coder (0.06), therapist (0.06) and condition (0.02). The Affective and Cognitive subscales had similar levels of variance attributable to youth (0.34 and 0.44 respectively) and residual error (0.40 and 0.35 respectively), followed by week (0.15), coder (0.10 and 0.03 respectively), condition (0.01 and 0.03 respectively) and therapist (0.00). For the Behavioral subscale, residual error accounted for half of the variance in coders' scores (0.50) followed by therapist (0.17), week (0.16), coder (0.04), and condition (0.00).

.

Chapter Seven

Discussion

This study reports on the development and initial score reliability and validity of the TPOCS-I, an in-session involvement instrument for youth therapy. Interrater reliability scores from trained graduate student coders who rated the TPOCS-I items were promising. The TPOCS-I scale and subscale scores evidenced convergent validity with an independent observer-rated measure of in-session involvement, providing evidence of convergent validity. However, the TPOCS-I scale and subscale scores demonstrated significant overlap with two observer-rated alliance measures, suggesting the TPOCS-I scores may not discriminate between alliance and involvement. Correlations between the TPOCS-I scale and subscales and other therapy process measures were in the expected direction and magnitude, providing some support for the construct validity of the measure. Lastly, the associations among the proposed TPOCS-I subscales two proposed subscale configurations (i.e., Behavioral, Affective, and Cognitive; Positive and Negative) did not perform as anticipated.

The current results generally indicate that the TPOCS-I can be coded reliably. Although some individual items yielded lower than anticipated interrater reliability scores across both samples (i.e., *Focus, Explores Feelings*), mean interrater reliability indicated "good" agreement for the Kendall sample and "excellent" agreement for the Child STEPS sample. In addition, the interrater reliabilities from the current study were comparable to that obtained by the CIRS, another observational measure of youth involvement in therapy (ICCs ranging from 0.07 to 0.78; Chu & Kendall, 2004). Despite some the concerns with individual item reliability, internal consistency of the TPOCS-I was also acceptable across samples and comparable to initial findings for the CIRS (alpha = 0.73; Chu & Kendall, 2004).

Non-normality in distribution was a concern for some items on the current measure (i.e., *Withdrawn, Distraction,* and *Oppositional*). These items were all part of the Negative subscale and assessed instances of behaviors demonstrating resistance or refusal to participate. In general, items forming the Negative involvement subscale were low occurring items and this may explain the non-normal distribution. It is possible that negative indictors of involvement, such as withdrawn or oppositional, may be low occurring behaviors in therapy. However, a larger, more diverse sample might be better suited for examining items such as those on the Negative involvement subscale. For example, a sample of youth with disruptive behavior problems might be more likely to see negative indicators of involvement.

One measurement issue this study was designed to address is the overlap between involvement and alliance. To address this question, the current study examined the extent to which the TPOCS-I overlapped with two observational instruments of alliance for youth therapy, the TPOCS-A and VTAS. The findings suggest that TPOCS-I total scores are not conceptually distinct from alliance. Similar effects were found between the CIRS and measures of the alliance indicating that the CIRS may also fail to demonstrate discriminant validity. These findings are consistent with previous studies examining the involvement-alliance association (Hudson et al., 2014; Karver et al., 2008; McLeod et al., 2013). One explanation for this overlap may be the source of measurement. Two different teams of expert coders were used when rating alliance and involvement. This methodology was used to reduce potential alliance-involvement association inflations due to using the same coders to rate each variable. Some have proposed the use of multiple perspectives when rating process measures such as involvement and alliance (Chu & Kendall, 2004; McLeod, Southam-Gerow, & Kendall, 2017). Perhaps the use of self-, parent-, or therapist-report of in-session involvement is better suited when attempting to distinguish alliance

and involvement and carve out the unique contributes of each to youth therapy (Chu & Kendall, 2004). However, no perspective is without methodological flaws; for instance youth may lack the cognitive development necessary to accurately rate involvement and ratings by parents and therapists may be influenced by perception of youth improvement (McLeod et al., 2017; Shirk & Saiz, 1992). In summary, the relationship between alliance and involvement is complex and intertwined and more research will be necessary to fully understand and detangle the unique boundaries and contributions of each variable within youth therapy.

Results provided mixed support for the validity of the TPOCS-I scale and subscale scores. The strong association between the TPOCS-I and the CIRS reveals conceptual overlap between the two measures. At first glance, this may appear to provide support for the construct validity of the TPOCS-I as both the TPOCS-I and CIRS are intended to measure in-session involvement. However, because neither the CIRS nor the TPOCS-I are distinguishable from the alliance measures, it is not clear that what variable these two measures actually assess. Thus, the findings do not provide clear support for the convergent validity of the TPOCS-I scores.

In addition, the association between the TPOCS-I and the CBAY-C, a measure of therapist competence, was in the "medium" range. Previous findings have examined therapist flexibility in intervention delivery, a proposed aspect of therapist competence (Chu & Kendall, 2009; Hudson et al., 2014), and the current findings are consistent with these previous studies. However, it should be noted that therapist flexibility and therapist competence are not synonymous and this limits definitive empirical comparison across studies. The association between involvement and competence in the current study is in line with the conceptual associations proposed by process researchers of child (Chu & Kendall, 2009) and adult therapy

(Cunha et al., 2012; Hill, 2005). More research will be necessary to establish baseline information regarding expected associations between involvement and competence.

The associations for between the TPOCS-I Total and subscale scores and a measure of theory-based therapeutic interventions, the TPOCS-RS, were also examined as an indicator of construct validity. It was expected that behavioral intervention-behavioral involvement, encouraging affect intervention-affective involvement, and cognitive intervention-cognitive involvement associations would fall in the "medium" range and all other intervention-involvement correlations would be in the "small" range. The findings from these analyses failed to consistently perform as anticipated. The distinction between behavioral, affective, and cognitive involvement may not have been clearly reflected in item generation and item grouping for each subscale. This may have impacted empirical results and specifically influenced the large overlap between subscales as demonstrated with correlation analyses between involvement subscales.

The findings regarding variance in TPOCS-I involvement score demonstrated that most variance was attributed to error. As expected, variance related to child accounted for the most non-error variance. This indicates youth vary in their level of involvement, which is consistent with previous research revealing an association between client factors and participation in the adult literature (Edelman & Chambless, 1993) and between parent factors and youth involvement (Nix et al, 2009). Some variance was also observed for time in treatment and this suggests that it may be important to measure in-session involvement over multiple sessions as opposed to adopting a more limited sampling plan. Therapist and coder effects appear to be minimal. However, the ability to detect these effects was likely impacted by the low level of nesting of clients within therapists, and the use of only two coders. Condition effects were not observed,

indicating that in-session involvement does not vary based on assigned treatment group for Child STEPS. This finding is consistent with expectations that involvement would not vary across treatment condition.

Another measurement issue this study was designed to address is the best way to assess involvement. In general, TPOCS-I Total and subscale scores were highly correlated and produced a similar pattern of correlations and this indicates that a TPOCS-I Total score may be most useful when assessing in-session involvement. The associations between the Behavioral, Affective, and Cognitive subscales did not perform as anticipated in validity analyses and due to the non-normal distribution of the Negative subscale, the use of the TPOCS-I Total score is advisable. To reduce coder burden, the 12-item Total score may be most appealing for future examination of in-session involvement. In addition, item revision to those items that displayed high overlap (i.e., *Initiate*, *Enthusiasm*, *Self-Disclose*, *Participation*) may improve overall scale performance and reduce redundancy between subscales.

In regards to the TPOCS-I subscales, the current study failed to provide clear indications regarding the use of either the Behavioral, Affective, and Cognitive or the Positive and Negative subscale configurations. Dividing involvement into behavioral, affective, and cognitive components may have conceptual traction (e.g., Morris et al., 2014), but this was not empirically supported in the current study. It is possible that item revision of the highly correlated items (e.g., providing distinguishing exemplars; combining items) on the TPOCS-I could reduce overlap between these three subscales. The use of the Positive and Negative subscales also lacked support as these scales demonstrated high overlap and the Negative subscale items occurred at a low frequency. The authors of the CIRS proposed the Positive and Negative domains, but these items were initially combined into a total score (Chu & Kendall, 2004).

Perhaps examination of these items with a larger or more diverse diagnostic sample would yield the opportunity to observe the instances of Negative involvement items and lead to greater opportunities to distinguish between these two subscales.

Although this study contains a number of strengths, there are also limitations to consider. First, the TPOCS-I is a measure intended to measure in-session involvement across common diagnostic and problem areas for youth. However, this sample consisted of a sample of youth with anxiety disorder diagnoses. This limits the generalizability of the current findings. It will be necessary to examine psychometric properties of the TPOCS-I within samples of youth across diagnostic areas to fully test the applicability of the TPOCS-I as a measure that is universal across youth therapies. Second, the current study is primarily correlational and therefore the data presented is not sufficient to make conclusions about the direction of effects. The current measure was intended to assess initial psychometric properties of the TPOCS-I, but did not examine the predictive validity of this measure. Previous conceptual and empirical research indicates that involvement follows the development of alliance and is a key variable in the prediction of client outcome. The current study was insufficient to examine these claims as outcome was not measured and data was correlational.

The current findings provide some indication of future directions when engaging in psychometric development of an in-session involvement measure. The initial psychometric properties of the TPOCS-I were not sufficient to definitively establish the construct validity of this measure scores. Previous research in the adult therapy literature indicates that involvement in therapy is related to client outcomes (Bohart & Tallman, 1999; Drieschner et al., 2004; Orlinsky et al., 1994). Involvement has been conceptually linked to outcome for the child literature (Doss, 2004) and there is some empirical evidence demonstrating the connection

between involvement and outcome for the youth literature (Karver et al., 2006). However, before any work can be done on the involvement-outcome connection and any implications resulting from such findings, one must be able to measure involvement. The current study demonstrates that the TPOCS-I can be reliably assessed, but it is unclear if this measure is a valid measure of client involvement. In fact, it is unclear if any of the current measures (e.g., CIRS) are valid measure of client involvement. Given the theoretical implications of an involvement-outcome association for youth therapy, it is arguably worthwhile to continue psychometric validation a measure of in-session involvement for youth. One next step for completing this task will be to examine the psychometric properties of this measure with a sample of youth who have diverse diagnostic presentations. Another next step will be to examine how this measure performs compared to ratings made from different perspectives. Last, item modification including revision, combination, and potentially addition may improve the performance of the TPOCS-I in regards to anticipated correlations with related measures.

In spite of methodological limitations, the current study provides initial psychometric data that supports the reliability of the TPOCS-I, an observational measure of in-session involvement for youth therapy. The support for the validity of the TPOCS-I is mixed and more research will be needed to establish this measure's convergent and discriminant validity. It is clear from the current findings that the association between the alliance and involvement is complex and thoughtful efforts will be necessary to untangle these two important but related variables.

Table 1.

Client and Therapist Descriptive Data by Study for Current Sample

	M (SD) or %	
Variable	CS Study	Kendall Study
Youth Level		
N	55	51
Age	9.89 *(1.71)*	10.35 (*1.89)*
Male	58.8	60.8
Race/ethnicity		
Caucasian	51.5	86.3
African-American	5.9	9.8
Latino	2.6	2.0
Asian-American/Pacific Islander	4.4	--
Multiracial	29.4	--
Other	2.9	2.0
Total no. diagnoses	2.95 (*2.01*)	3.02 (*1.45*)
CBCL T-score		
Total	65.13 (*7.81*)	63.18 (*8.44*)
Internalizing	69.51 (*7.18*)	67.40 (*8.37*)
Externalizing	58.22 (*10.69*)	52.96 (*10.08*)
BIS		NC
Total	16.02 (*8.21*)	
Interpersonal	5.16 (*3.95*)	
School/Work	5.96 (*3.57*)	
Self-fulfillment	5.57 (*2.85*)	
Family Level		
Age	40.53 (*9.76*)	NC
Male	29.1	NC
Annual family income (< $60K)	57.3	35.3
Annual family income (> $60K)	36.8	56.9
Therapist Level		
N	35	16
Age	40.27 (*10.21*)	NC
Male	19.2	21.8
Ethnicity/race		
Caucasian	44.4	68.6
African-American	5.0	--
Latino/Hispanic	1.6	17.6
Asian-American/Pacific Islander	22.6	9.8
Multiracial	8.1	--
Degree type		NC
MA/MSW	37.8	
PhD	8.4	
PsyD	8.4	
MD	0.8	
EdD	0.8	
LCSW	11.8	
MFCC/MFT	5.0	

Other	47.9	
Years of experience	5.75 (*7.23*)	NC
Principal theoretical orientation		NC
CB/C/B	37.4	
Eclectic	9.8	
Family systems	5.7	
Psychodynamic	6.6	
Other	33.0	

Note. Chars = characteristics; NC = not collected; CS = modular and standard individual cognitive behavioral therapy conditions of Child STEPs study (Weisz et al., 2012); Kendall Study = ICBT of Kendall et al. (2008) study. CBCL = Child Behavior Checklist; CB = Cognitive Behavioral; C = Cognitive; B = Behavioral.

Table 2

TPOCS-I Item Description, Sources, and Proposed Subscale

Item Description	SS	New item	CIRS	SCCS	CPPS	VPPS	AOR	TPOCS-A	VTAS-R
5. Focus: Does the client remain focused on a specific therapeutic activity?	B,P					x			
8. Behavioral Involvement: Does the client engage in therapeutic activities that require non-verbal involvement?	B,P	x							
9. Participation: Does the client participate in therapeutic activities?	B,P			x	x	x			
12. Distraction: Does the client attempt to distract the activities away from therapy-related activities (verbally or physically)?	B,N		x	x			x		
13. Oppositional: Is the client oppositional to therapeutic activities?	B,N		x		x	x	x	x	x
14. Passive: Does the client rely on the therapist to determine therapeutic activities?	B,N					x			
2. Demonstrate Enthusiasm: Does the client demonstrate enthusiasm when engaged in therapeutic activities?	A,P		x		x	x			
6. Explores Feelings: Does the client explore feelings in response to therapeutic activities?	A,P				x	x			
11. Inhibited/Avoidant: Is the client inhibited or avoidant of therapeutic activities (e.g., not fully participating)?	A,N		x					x	
1. Initiate Discussion: Does the client initiate discussions or introduce new topics of conversation related to the therapeutic activities?	C,P		x			x			
3. Self-Disclosure: Does the client offer personal information (self-disclose) when related to therapeutic activities?	C,P		x					x	x
4. Ask Questions: Does the client ask questions and/or for further explanation about therapeutic activities?	C,P		x	x		x			

7. Cognitive Involvement: Does the client explore cognitions/ perceptions in response to therapeutic activities?	C,P	x			
10. Withdrawn/Disinterested: Does the client withdraw from therapeutic activities (e.g., not respond to the therapist)?	C,P		x		x

Note. The CIRS is from Chu and Kendall (2004); the SCCS is from Karver and colleagues (2008); the CPPS is from Estrada and Russell (1999), the VPPS is from Gomes-Schwartz (1978); the AOR is from Karver and colleagues (2008); the TPOCS-A is from McLeod and Weisz (2005); and the VTAS is from Hartley and Strupp (1983).

Table 3

Therapy Process Observational Coding Scale – Involvement (TPOCS-I): Descriptive data and reliability results Kendall sample

Item	Brief Description	Range	M	SD	ICC	Skewness M	Kurtosis M
Initiate	Initiate discussions or introduce new topics of conversation	5	2.66	0.94	0.58	-0.37	0.14
Enthusiasm	Demonstrate enthusiasm when engaged in therapeutic activities	4.5	2.99	0.89	0.67	-0.94	1.25
Self-Disclose	Offer personal information (self-disclose)	4	3.08	0.92	0.59	-0.19	-0.44
Ask Questions	Ask questions and/or for further explanation about therapeutic activities	5	2.21	1.09	0.49	-0.05	-0.38
Focus	Remain focused on a specific therapeutic activity	4.5	3.41	0.86	0.52	-0.67	0.73
Explores Feelings	Explore feelings in response to therapeutic activities	4	2.76	0.87	0.45	-0.32	0.04
Cognitive	Explore cognitions/perceptions	4.5	2.91	1.09	0.66	-0.30	-0.69
Behavioral	Engage in therapeutic activities that require non-verbal involvement	4.5	3.52	1.08	0.70	-1.13	0.61
Participation	Participate in therapeutic activities	4	3.51	0.76	0.67	-1.10	2.38
Withdrawn	Withdraw or not respond to the therapist	4	4.8	0.61	0.86	-4.74	25.12
Inhibited	Inhibited or avoidant of therapeutic activities (e.g., not fully participating)	2.5	4.49	0.57	0.16	-1.15	1.11
Distraction	Attempt to distract the activities from therapeutic activities verbally or physically	5	4.59	0.81	0.65	-3.34	13.94
Oppositional	Oppositional to therapeutic activities	4	4.87	0.48	0.72	-6.18	46.56
Passive	Rely on the therapist to determine therapeutic activities	3.5	3.76	0.83	0.53	-0.70	0.17
Total		-	-	-	0.73	-	-

Note. $N = 95$. Interrater reliability based on model ICC (2,2).

Table 4

Therapy Process Observational Coding Scale – Involvement (TPOCS-I): Descriptive data and reliability results Child Steps sample

Item	Brief Description	Range	M	SD	ICC	Skewness M	Kurtosis M
Initiate	Initiate discussions or introduce new topics of conversation	5	2.45	1.12	0.73	-0.24	-0.31
Enthusiasm	Demonstrate enthusiasm when engaged in therapeutic activities	5	2.71	1.04	0.74	-0.51	-0.08
Self-Disclose	Offer personal information (self-disclose)	5	2.78	1.05	0.74	-0.23	-0.43
Ask Questions	Ask questions and/or for further explanation about therapeutic activities	5	1.92	1.19	0.64	-0.01	-0.75
Focus	Remain focused on a specific therapeutic activity	5	3.15	0.82	0.49	-0.27	-0.12
Explores Feelings	Explore feelings in response to therapeutic activities	5	2.61	0.92	0.56	-0.26	-0.10
Cognitive	Explore cognitions/perceptions	5	2.25	1.08	0.65	-0.07	-0.70
Behavioral Participation	Engage in therapeutic activities that require non-verbal involvement	5	2.46	1.32	0.62	-0.22	-0.98
Withdrawn	Participate in therapeutic activities	4.5	3.07	0.82	0.68	-0.47	-0.05
Inhibited	Withdraw or not respond to the therapist	5	4.6	0.89	0.78	-2.75	7.54
Distraction	Inhibited or avoidant of therapeutic activities (e.g., not fully participating)	4.5	4.50	0.84	0.69	-2.20	4.90
Oppositional	Attempt to distract the activities from therapeutic activities verbally or physically	4.5	4.69	0.64	0.56	-2.96	10.63
Passive	Oppositional to therapeutic activities	4	4.84	0.53	0.73	-4.48	23.38
Total	Rely on the therapist to determine therapeutic activities	5	3.52	1.06	0.73	-0.77	0.26
		-	-	-	0.82	-	-

Note. N = 734. Interrater reliability based on model ICC (2,2).

Table 5.

Correlations among items of the Therapy Process Observational Coding Scale – Involvement (TPOCS-I).

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.
Behavioral														
1. Focus	--	.15**	.61**	.56**	.30**	.08*	.40**	.48**	.24**	.27**	.41**	.17**	.44**	.26**
2. Behavioral	.12	--	.28**	-.01	.18**	.15**	.30**	.15**	.09*	.17**	.09*	.20**	.06	.20**
3. Participation	.61**	.34**	--	.16**	.30**	.57**	.78**	.64**	.45**	.71**	.74**	.57**	.60**	.60**
4. Distract[a]	.66**	.09	.30**	--	.26**[b]	0.23**[b]	.06	.12**	.03[b]	-.06	.04	-.20**	.07	-.09**[b]
5. Oppositional[a]	.39**	.38**	.45**	.06[b]	--	-.15**[b]	.33**	.19**	.19**[b]	.08*	.16**	.01	.11**	.09**[b]
6. Passive[a]	.01	.03	.51**	-.20**[b]	.01[b]	--	.57**	.36**	.23**[b]	.72**	.61**	.61**	.43**	.44**[b]
Affective														
7. Demonstrate	.28**	.33**	.73**	.05	.45**	.44**	--	.43**	.55**	.72**	.67**	.59**	.46**	.64**
8. Explores Feelings	.31**	.07	.52**	.23*	.16	.23*	.34**	--	.23**	.48**	.57**	.42**	.53**	.35**
9. Inhibited[a]	.13	.25*	.48**	-.22*[b]	.23*[b]	.37**[b]	.61**	.18	--	.48**	.46**	.34**	.36**	.46**
Cognitive														
10. Initiate	.14	.14	.62**	-.08	.26*	.58**	.71**	.38**	.51**	--	.81**	.68**	.53**	.60**
11. Self-Disclose	.33**	.07	.68**	.05	.26*	.45**	.64**	.60**	.38**	.68**	--	.59**	.59**	.57**
12. Ask Questions	.01	.08	.49**	-.09	.12	.57**	.49**	.20	.34**	.67**	.49**	--	.45**	.52**
13. Cognitive	.40**	.14	.53**	.23*	.14	.31**	.33**	.65**	.16	.35**	.60**	.19	--	.45**
14. Withdrawn[a]	.12	.34**	.55**	-.14[b]	.26**[b]	.38**[b]	.57**	.32	.41[b]	.44**	.38**	.40**	.32**	--

Note. Kendall sample (N = *95)* data is below the diagonal and Child STEPS (N = 734) is above the diagonal.
[a] Negatively worded items were reverse scored before calculating correlations; [b] Correlations for these items were calculated using Kendall's tau due to their skewed and kurtotic distribution.
*p <0.05 level (2 tailed); **p<0.01 level (2 tailed)

Table 6.

Correlations among items of the Therapy Process Observational Coding Scale – Involvement (TPOCS-I).

	1.	2.	3.	4.	5.	6.	7.	8.	9.	.10	.11	.12	.13	.14
Positive														
1. Initiate	-	.72**	.81**	.68**	.27**	.48**	.53**	.17**	.71**	.60**	.48**	-.06	.08*	.72**
2. Enthusiasm	.71**	-	.67**	.59**	.40**	.43**	.46**	.30**	.78**	.64**	.55**	.06	.33**	.57**
3. Self-Disclose	.68**	.64**	-	.59**	.41**	.57**	.59**	.09*	.74**	.57**	.46**	.04	.16**	.61**
4. Ask Questions	.67**	.49**	.49**	-	.17**	.42**	.45**	.20**	.57**	.52**	.34**	-.20	.01	.61**
5. Focus	.14	.28**	.33**	.01	-	.48**	.44**	.15**	.61**	.26**	.24**	.56**	.30**	.08*
6. Explores Feelings	.38**	.34**	.60**	.19	.31**	-	.53**	.15**	.64**	.35**	.23**	.12**	.19**	.36**
7. Cognitive	.35**	.33**	.60**	.19	.40**	.65**	-	.06	.60**	.45**	.36**	.07	.11**	.43**
8. Behavioral	.14	.33**	.07	.08	.12	.07	-.14	-	.38**	.20**	.09*	-.01	.18**	.15**
9. Participation	.62**	.73**	.68**	.49**	.61**	.52**	.53**	.34**	-	.60**[b]	.48[b]	.16**[b]	.30**[b]	.57**
Negative														
10. Withdrawn[a]	.44**	.57**	.38**	.40**	.12	.32**	.32**	.34**	.55**	-	.46**[b]	-.09**[b]	.09**[b]	.64**
11. Inhibited[a]	.51**	.60**	.38**	.34**	.13	.18	.16	.25*	.48**	.41**[b]	-	.03[b]	.19**[b]	.42**
12. Distraction[a]	-.08	.05	.05	-.09	.66**	.23*	.23*	.09	.30**	-.14[b]	-.22*[b]	-	.26**[b]	-.23**
13. Oppositional[a]	.26*	.45**	.26*	.12	.39**	.16	.14	.38**	.45**	.26**[b]	.23*[b]	.06[b]	-	-.17**
14. Passive[a]	.58**	.44**	.45**	.57**	.01	.23*	.31**	.03	.51**	.54**	.42**	-.22*	-.05	-

Note. Kendall sample (N = *95)* data is below the diagonal and Child STEPS (N = 734) is above the diagonal.
[a] Negatively worded items were reverse scored before calculating correlations; [b] Correlations for these items were calculated using Kendall's tau due to their skewed and kurtotic distribution.
*p <0.05 level (2 tailed); **p<0.01 level (2 tailed)

Table 7

Therapy Process Observational Coding Scale – Involvement (TPOCS-I): Descriptive subscale data and internal consistency Kendall sample

Scale/Subscale	Brief Description	Range	*M*	*SD*	ICC	Alpha	Skewness *M*	Kurtosis *M*
Total	14 item total score	2.93	3.54	0.73	0.73	0.86	-1.56	4.27
Total-12	12 item total score	3.25	3.34	0.57	0.73	0.87	-1.36	3.42
Behavioral	6 item subscale (i.e., *Focus*, *Behavioral*, *Participation*, *Positive*, *Distraction*, *Oppositional*, *Passive*)	3.33	3.94	0.50	0.69	0.65	-2.02	7.01
Behavioral-4	4 item subscale (i.e., *Focus*, *Behavioral*, *Participation*, *Positive*, *Passive*)	3.00	3.55	0.58	0.69	0.56	-1.27	2.32
Affective	3 item subscale (i.e., *Demonstrate*, *Explores feelings*, *Inhibited*)	3.17	3.41	0.60	0.57	0.62	-1.04	2.26
Cognitive	5 item subscale (i.e., *Initiate*, *Self-disclose*, *Ask*, *Cognitive*, *Withdrawn*)	4.10	3.13	0.70	0.72	0.79	-0.93	2.12
Positive	9 item subscale	3.88	3.38	0.71	0.69	0.84	-1.02	2.16
Negative	5 item subscale (i.e., *Withdrawn*, *Inhibited*, *Distraction*, *Oppositional*, *Passive*)	2.40	4.50	0.40	0.70	0.52	-2.44	8.49
Negative -3	3 item subscale (i.e., *Withdrawn*, *Inhibited*, *Passive*)	3.00	4.35	0.54	0.71	0.71	-1.93	5.59

Note. $N = 95$.

Table 8

Therapy Process Observational Coding Scale – Involvement (TPOCS-I): Descriptive subscale data and internal consistency Child STEPS sample

Scale/Subscale	Brief Description	Range	*M*	*SD*	ICC	Alpha	Skewness *M*	Kurtosis *M*
Total	14 item total score	3.57	3.25	0.62	0.80	0.89	-0.83	0.55
Total-12	12 item total score	4.08	3.00	0.71	0.81	0.90	-0.87	0.61
Behavioral	6 item subscale (i.e., *Focus*, *Behavioral*, *Participation*, *Positive*, *Distraction*, *Oppositional*, *Passive*)	3.50	3.62	0.52	0.68	0.59	-0.48	0.48
Behavioral-4	4 item subscale (i.e., *Focus*, *Behavioral*, *Participation*, *Positive*, *Passive*)	4.00	3.05	0.69	0.67	0.61	-0.37	0.16
Affective	3 item subscale (i.e., *Demonstrate*, *Explores feelings*, *Inhibited*)	4.33	3.27	0.73	0.75	0.67	-0.89	0.73
Cognitive	5 item subscale (i.e., *Initiate*, *Self-disclose*, *Ask*, *Cognitive*, *Withdrawn*)	4.50	2.80	0.87	0.82	0.87	-0.81	0.33
Positive	9 item subscale	4.50	2.92	0.84	0.74	0.88	-0.46	-0.21
Negative	5 item subscale (i.e., *Withdrawn*, *Inhibited*, *Distraction*, *Oppositional*, *Passive*)	2.90	4.43	0.50	0.81	0.59	-1.91	4.00
Negative -3	3 item subscale (i.e., *Withdrawn*, *Inhibited*, *Passive*)	4.67	4.29	0.78	0.84	0.79	-02.11	4.91

Note. $N = 734$.

Table 9.
Correlations among subscales of the Therapy Process Observational Coding Scale – Involvement (TPOCS-I).

	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. Total	-	.99**	.86**	.87**	.91**	.93**	.98**	.86**	.81**
2. Total – 12	.99**	-	.82**	.87**	.91**	.95**	.98**	.84**	.83**
3. Behavioral	.83**	.77**	-	.95**	.72**	.64**	.84**	.72**	.57**
4. Behavioral – 4	.85**	.84**	.94**	-	.72**	.71**	.88**	.69**	.66**
5. Affective	.89**	.89**	.64**	.67**	-	.82**	.88**.	.84**	.79**
6. Cognitive	.90**	.93**	.54**	.62**	.79**	-	.92**	.78**	.81**
7. Positive	.98**	.98**	.78**	.82**	.88**	.92**	-	.74**	.71**
8. Negative	.85**	.80**	.82**	.77**	.75**	.67**	.73**	-	.92**
9. Negative – 3	.74**	.78**	.52**	.66**	.70**	.73**	.66**	.82**	-

Note. Kendall sample (N = *95)* data is below the diagonal and Child STEPS (N = 734) is above the diagonal.
*p <0.05 level (2 tailed); **p<0.01 level (2 tailed)

Table 10.
Correlations among the subscales of the Therapy Process Observational Coding Scale – Involvement (TPOCS-I), and other measures of therapy process

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1. Total - 12	-	.87**	.91**	.95**	.98**	.83**	-	.73**	.76**	.28**	.09*	.12**	.28**
2. Behavioral - 4	.84**	-	.72**	.71**	.88**	.66**	-	.62**	.65**	.25**	.22**	.10**	.17**
3. Affective	.89**	.67**	-	.82**	.88**	.79**	-	.73**	.71**	.28**	.06	.11**	.28**
4. Cognitive	.93**	.62**	.79**	-	.92**	.81**	-	.68**	.72**	.25**	.01	.12**	.31**
5. Positive	.98**	.82**	.88**	.92**	-	.71**	-	.72**	.75**	.31**	.15**	.13**	.31**
6. Negative - 3	.78**	.66**	.70**	.73**	.66**	-	-	.60**	.62**	.13**	-.08*	.07	.16**
7. CIRS	.72**	.54**	.69**	..69**	.66**	.72**	-	-	-	-	-	-	-
8. TPOCS-A	.62**	.54**	.67**	.51**	.57**	.64**	.68**	-	.82**	.34**	.13**	.07	.30**
9. VTAS-R	.72**	.63**	.67**	.63**	.70**	.59**	.74**	.69**	-	.38**	.10**	.13**	.35**
10. CBAY-C	.26*	.25*	.31**	.18	.24*	.24*	.25*	.39**	.40**	-	.47**	-.12**	.62**
11. TPOCS-RS Bx	.03	.24*	.03	-.11	.04	.02	-.07	.12	.07	.38**	-	-.15**	.13**
12. TPOCS-RS EA	.19*	.08	.19	.21*	.22*	.03	.07	.08	.13	-.04	.03	-	-.02
13. TPOCS-RS Cog	.28**	.02	.28**	.39**	.29**	.15	.34**	.15	.25*	.22*	-.13	.26**	-

Note. Kendall sample (N = *95)* data is below the diagonal and Child STEPS (N = 734) is above the diagonal. CIRS correlations were not coded for the Child STEPS sample and are therefore not displayed.
*p <0.05 level (2 tailed); **p<0.01 level (2 tailed)

Table 11. *Variance components for TPOCS-I*

		Variance Components					
		Coder	Condition	Therapist	Child	Week	Residual
Kendall sample	Total 12	0.08	-	0.07	0.33	0.11	0.39
N = 95	Behavioral	0.07	-	0.05	0.31	0.09	0.48
	Affective	0.10	-	0.08	0.25	0.02	0.55
	Cognitive	0.05	-	0.04	0.28	0.22	0.43
Child STEPS sample	Total 12	0.06	0.02	0.06	0.36	0.15	0.35
N = 734	Behavioral	0.04	0.00	0.17	0.13	0.16	0.50
	Affective	0.10	0.01	0.00	0.34	0.15	0.40
	Cognitive	0.03	0.03	0.00	0.44	0.15	0.35

Note: Variance components estimates represent the portion of variance that is attribute to each source of variance

References

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA School-Age Forms and Profiles.* Burlington: University of Vermont, Center for Children, Youth and Families.

Adelman, H., Kaser-Boyd, N., & Taylor, L. (1984). Children's participation in consent for psychotherapy and their subsequent response to treatment. *Journal of Clinical Child Psychology, 11*(2), 170-178.

Bakeman, R., & Quera, V. (2012). Behavioral observation. In H. Cooper, P. Camic, D. Long, A. Panter, D. Rindskopf, & K. Sher (Eds.), *APA Handbook of Research Methods in Psychology* (Vol. 1, pp. 207-225). Washington D.C.: American Psychological Association.

Barber, J.P., Foltz, C., Crits-Christoph, P., & Chittams, J. (2004). Therapists' adherence and competence and treatment discrimination in the NIDA collaborative cocaine treatment study. *Journal of Clinical Psychology, 60,* 29-41.

Barkley, R. (1997). *Defiant children: A clinician's manual for assessment and parent training* (2nd ed.). New York, New York: Guilford Press.

Becker, K., Kiser, L., Herr, S., Stapelton, L., Barksdale, C., & Buckingham, S. (2014). Changes in treatment engagement of youth and families with complex needs. *Children and Youth Services Review, 46*, 276-284.

Bellg, A., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D., Ory, M., . . . Czajkowski, S. (2004). Enhancing treatment fidelity in health behavior change studies: Practices and

recommendations from the NIH Behavior Change Consortium. *Health Psychology, 23*(5), 443-451.

Beutler, L. (2014). Welcome to the party, but…. *Psychotherapy, 51*(4), 496-499.

Beutler, L., Moliero, C., & Talbi, H. (2002). How practitioners can systematically use empirical evidence in treatment selection. *Journal of Clinical Psychology, 58*(10), 1199-1212.

Bohart, A. & Tallman, K. (1999). What do we mean by the client as active self healer?. In A. Bohart & K. Tallman (Eds). *How Clients Make Therapy Work: The Process of Active Self-Healing*. (pp. 3-23). Washington D.C.: American Psychological Association.

Burgio, L., Corcoran, M., Lichstein, K., Nichols, L., Czaja, S. Gallagher-Thompson, D. …for the REACH Investigators. (2001). Judging outcomes in psychosocial interventions for dementia caregivers: The problem of treatment implementation. *Gerontologist, 41*(4), 481-489.

Burns, D. D., & Nolen-Hoeksema, S. (1991). Coping styles, homework compliance, and the effectiveness of cognitive-behavioral therapy. *Journal of Consulting and Clinical Psychology, 59*(2), 305-311.

Carlson, J., Watts, R., & Maniacci, M. (2006). *Adlerian Therapy: Theory and Practice.* Washington, DC: American Psychological Association.

Chambers, C. & Johnston, C. (2002). Developmental differences in children's use of rating scales. *Journal of Pediatric Psychology, 27*(1), 27-36.

Chiu, A. W., McLeod, B. D., Har, K., & Wood, J. J. (2008). Child-therapist alliance and clinical outcome sin cognitive behavioral therapy for child anxiety disorders. *Journal of Child Psychiatry and Psychology, 50*(6), 751-758.

Chorpita, B., Daleiden, E., & Weisz, J. (2005). Identifying and selecting the common elements of evidence based interventions: A distillation and matching model. *Mental Health Services Research*, *7*(1), 5-20.

Chorpita, B. F. & Weisz, J. R. (2005). *Modular approach to therapy for children with anxiety, depression, or conduct problems*. Honolulu, HI: University of Hawaii at Manoa; Boston, MA: Judge Baker Children's Center, Harvard Medical School

Chu, B. C., & Kendall, P. C. (1999). *Child Involvement Rating Scale (CIRS) Scoring Manual*. Unpublished coding manual.

Chu, B. C. & Kendall, P. C. (2004). Positive association of child involvement and treatment outcome with a manual-based cognitive-behavioral treatment for children with anxiety. *Journal of Consulting and Clinical Psychology*, *72*(5), 821-829.

Chu, B. C. & Kendall, P. C. (2009). Therapist responsiveness to child engagement: Flexibility within manual-based CBT for anxious youth. *Journal of Clinical Psychology*, *65*(7), 736-754.

Chu, B., Suveg, C., Creed, T., & Kendall, P. (2010). Involvement shifts, alliance ruptures, and management engagement over therapy. In D. Castro-Blanco & M. Karver (Eds.) *Elusive alliance: Treatment engagement strategies with high-risk adolescents* (pp. 95-121). Washington DC: American Psychological Association.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*, 284–290.

Clark, L. A., & Watson, D. (2003). Constructing validity: Basic issues in objective scale development. In A. Kazden (Ed.) *Methodological issues & strategies in clinical research* (3rd ed.)., (pp. 207-231). Washington, DC: American Psychological Association.

Clarke, A., Marshall, S., Mautone, J., Soffer, S., Jones, H. A., Costingan, T.,… Power, T. J. (2013). Parent attendance and homework adherence predict response to a family-school intervention for children with ADHD. *Journal of Clinical Child and Adolescent Psychology. 44* (1), 58-67.

Creed, T. & Kendall, P. (2005). Therapist alliance-building behavior within cognitive-behavior treatment for anxiety in youth. *Journal of Consulting and Clinical Psychology, 73*(3), 498-505.

Cunha, C., Goncalves, M. M., Hill, C. E., Mendes, I., Ribeiero, A. P., Sousa, I.…Greenberg, L.S. (2012). Therapist interventions and client innovative moments in emotion-focused therapy for depression. *Psychotherapy, 49*(4), 536-548.

De Araujo, L., Ito, L., & Marks, I. (1996). Early compliance and other factors predicting outcome of exposure of obsessive-compulsive disorder. *British Journal of Psychiatry, 169*, 747-752.

De Jong, P. & Berg, I. (2008). *Interviewing for solutions* (3rd ed.). Belmont, CA: Brooks/Cole.

DeVellis, R. F. (2003). *Scale Development: Theory and Application* (2nd ed.). Thousand Oaks, California: Sage Publications, Inc.

Diamond, G. M., Liddle, H. A., Hogue, A., & Dakof, G. A. (1999). Alliance-building interventions with adolescents in family therapy: A process study. *Psychotherapy, 36*, 355-368.

Doss, B. D. (2004). Changing the way we study change in psychotherapy. *Clinical Psychology: Science and Practice*, *11*, 368-386.

Drieschner, K., Lammers, S., & van der Staak, C. (2004). Treatment motivation: An attempt for clarification of an ambiguous concept. *Clinical Psychology Review*, *(23)*, 1115-1137.

Ebesutani, C., Berstein, A., Martinez, J. I., & Weisz, J. R. (2011). The Youth Self-Report: Applicability and validity across younger and older youths. *Journal of Clinical Child and Adolescent Psychology*, *40*, 338-346.

Edelman, R. E., & Chambless, D. L. (1993). Compliance during sessions and homework in exposure-based treatment of agoraphobia. *Behavior Respiratory Therapy*, *31*(8), 767-773.

Ellis, A. (1994). *Reason and emotion in psychotherapy revised.* New York: Kensington.

Ellis, A. (1996). *Better, deeper, and more enduring brief therapy: The rational emotive behavior therapy approach.* New York: Brunner/Mazel.

Estrada, A., & Russell, R. (1999). The development of the Child Psychotherapy Process Scales (CPPS). *Psychotherapy Research*, *9*, 154-166.

Fayers, P. M., Jordhoy, M. S., & Kaasa, S. (2002). Cluster-randomized trials. *Palliative Medicine*, *16, 69-70*.

Fjermestad, K. W., McLeod, B. D., Tully, C. B., & Liber, J. M. (2015). *Therapist characteristics and interventions: Enhancing alliance and involvement with youth.* Manuscript in preparation.

Flannery-Schroeder, E. C., & Kendall, P. C. (2000). Group and individual cognitive-behavioral treatments for youth with anxiety disorders: A randomized clinical trial. *Cognitive Therapy and Research*, *24*, 251-278.

Flicker, S. M., Turner, C. W., Waldron, H. B., Brody, J. L., & Ozechowski, T. J. (2008). Ethnic background, therapeutic alliance, and treatment retention in functional family therapy with adolescents who abuse substances. *Journal of Family Psychology, 22*(1), 167-170.

Fristad, M. A., Glickman, A. R., Verducci, J. S., Teare, M., Weller, E. B., & Weller, R. A. (1998). Children's interview for psychiatric syndromes (ChIPS): Psychometrics in two community samples. *Journal of Child and Adolescent Pharmacology*, *8*(4), 237-245.

Garland, A. F., Haine-Schlagel, R., Accurson, E. C., Baker-Ericzen, M. J., & Bookman-Frazee, L. (2012). Exploring the effect of therapists' treatment practices on client attendance in community-based care for children. *Psychological Services*, *9*(1), 74-88.

Glasser, W. (1992). Reality therapy. *New York State Journal for Counseling and Development*, *7*(1), 5-13.

Gomes-Schwartz, B. (1978). Effective ingredients in psychotherapy: Prediction of outcome from process variables. *Journal of Consulting and Clinical Psychology*, *46*(5) 1023-1035.

Gorin, S. S. (1993). The prediction of child psychotherapy outcome: Factors specific to treatment. *Psychotherapy*, *30*(1), 152-158.

Hartley, D. E. & Strupp, H. H. (1983). The therapeutic alliance: Its relationship to outcome in brief psychotherapy, In M. J. Hillsdale (Ed.), *Empirical Studies of Psychoanalytic Theories* (pp. 7-11), New Jersey: Analytic Press.

Herbert, J. R., Ma, Y., Ebbeling, C. B., Matthews, C. E., & Ockene, I. S. (2001). Self-report data. In L. Burke & I. Ockene (Eds.) *Compliance in Healthcare and Research* (pp. 163-179). Armonk, NY: Futura Publishing Company, Inc.

Hill, C. E. (2005). Therapist techniques, client involvement, and the therapeutic relationship: Inextricably intertwined in the therapy process. *Psychotherapy: Theory, Research, Practice, Training, 42*(4), 431-442.

Hill, C. E., Crook-Lyon, R. E., Hess, S. A., Goates-Jones, M., Roffman, M. Stahl, J., … Johnson, M. (2006). Prediction of session process and outcome in the Hill dream model: Contributions of client characteristics and the process of the three stages. *Dreaming, 16*(3), 159-185.

Hogue, A., Liddle, H. A., & Rowe, C. (1996). Treatment adherence process research in family therapy: A rational and some practical guidelines. *Psychotherapy, 33*(2), 332-345.

Horvath, A. O. & Greenberg, L. S. (1989). Development and validation of the working alliance inventory. *Journal of Counseling Psychology*, *36*(2), 223-233.

Howard, B., Chu, B.C., Krain, A. L., Marrs-Garcia, M. A., & Hedtke, K. A. (2004). *Cognitive-behavioral family therapy for anxious children: Therapist manual* (2nd ed.). Ardmore, PA: Workbook.

Hudson, J. L., Kendall, P. C., Chu, B. C., Gosch, E., Martin, E., Taylor, A., & Knight, A. (2014). Child involvement, alliance, and therapist flexibility: Process variables in cognitive-behavioural therapy for anxiety disorders in childhood. *Behaviour Research and Therapy, 52*, 1-8.

Jungbluth, H. J. & Shirk, S. R. (2009). Strategies for building involvement in cognitive-behavioral therapy for adolescent depression. *Journal of Consulting and Clinical Psychology, 77*(6), 1179-1184.

Kamin, D. J., Garske, J. P., Sawyer, P. K., & Rawson, J. C. (1993). Effects of explicit time-limits on the initial therapeutic alliance. *Psychological Reports*, *72*(2), 443-448.

Karver, M. S., Handelsman, J. B., Fields, S., & Bickman, L. (2006). Meta-analysis of therapeutic relationship variable in youth and family therapy: The evidence for different relationship variables in the child and adolescent treatment literature. *Clinical Psychology Review, 26*, 50-65.

Karver, M. S., Handelsman, J. B., Fields, S., & Bickman, L. (2005). A theoretical model of common process factors in youth and family therapy. *Mental Health Services Research, 7*(1), 35-51.

Karver, M., Shirk, S., Handelsman, J. B., Fields, S., Crisp, H., Gudmundsen, G., McMakin, D. (2008). Relationship processes in youth psychotherapy: Measuring alliance, alliance-building behaviors, and client involvement. *Journal of Emotional and Behavioral Disorders, 16*(1), 15-28.

Kazantzis, N., Deane, F. P., & Ronan, K. R. (2000). Homework assignments in cognitive and behavioral therapy: A meta-analysis. *Clinical Psychology: Science and Practice, 7*(2), 189-202.

Kendall P.C. (1994). Treating anxiety disorders in children: Results of a randomized clinical trial. *Journal of Consulting and Clinical Psycho*logy, *62*(1): 100-110.

Kendall P.C. & Hedtke, K.A. (2006a). *Cognitive-behavioral therapy for anxious children: Therapist manual* (3rd ed.). Ardmore, PA: Workbook.

Kendall P.C. & Hedtke, K.A. (2006b). *Coping cat workbook* (2nd ed.). Ardmore, PA: Workbook.

Kendall, P. C., Flannery-Schroeder, E., Panichelli-Mendel, S. M., Southam-Gerow, M., Henin, A., & Warman, M. (1997). Therapy for youths with anxiety disorders: A second randomized clinical trial. *Journal of Consulting and Clinical Psychology, 65*, 366-380.

Kendall, P. C., Hudson, J. L, Gosch, E., Flannery-Schroeder, E. F., & Suveg, C. (2008).

Cognitive Behavioral Therapy for anxiety disordered youth: A randomized clinical trial evaluating child and family modalities. *Journal of Consulting and Clinical Psychology, 76*, 282-297.

Kendall, P. C., Kane, M., Howard, B., & Siqueland, L. (1990). *Cognitive behavioral treatment of anxious children: Treatment manual*. Ardmore, PA: Workbook.

Krain, A., Hudson, J., Choudhury, M., & Kendall, P. C. (2000). Family education, support, and attention for child anxiety. Unpublished therapist manual, Temple University.

Krupnick, J. L., Sotsky, S. M., Simmens, S., Moyer, J., Elkin, I., Watkins, J., & Pilkonis, P. A. (1996). The role of the therapeutic alliance in psychotherapy and pharmacotherapy outcome: Findings in the National Institute of Mental Health treatment of depression collaborative research program. *Journal of Consulting and Clinical Psychology*, *64*, 532-539.

Lambert, M. J. (2010). Therapist effects. In M. J. Lambert (Ed), *Prevention of treatment failure: The use of measuring, monitoring, and feedback in clinical practice* (pp. 175-202). Washington, DC: American Psychological Association.

Lambert, M., & Hill, C. (1994). Assessing psychotherapy outcomes and processes. In A. Bergin and S. Garfield (Eds.), *Handbook of psychotherapy and behavior change (4th ed.,* pp. 72-113), New York: John Wiley.

Lerner, M. D., Mikami, A. Y., & McLeod, B. D., (2011). The alliance in a friendship coaching intervention for parents of children with ADHD. *Behavior Therapy*, *42*, 449-461.

Liber, J. M., McLeod, B. D., Van Widenfelt, B. M., Goedhart, A. W., van der Leeden, A. J. M., Utens, E. M. W. J., … Treffers, P. D. A. (2010). Examining the relation between the

therapeutic alliance, treatment adherence, and outcome of cognitive behavioral therapy for children with anxiety disorders. *Behavior Therapy*, *41*, 172-186.

Lichstein, K. L., Reidel, G. W., & Grieve, R. (1994). Fair tests of clinical trials: A treatment implementation model. *Advances in Behaviour Research & Therapy*, *16*(1), 1-20.

Lindheim, O. & Kolko, D. J. (2010). Trajectories of symptom reduction and engagement during treatment for childhood behavior disorders: Differences across settings. *Journal of Abnormal Child Psychology, 38*, 995-1005.

Littell, J. H. (2001). Client participation and outcomes of intensive family preservation services. *Social Work Research*, *25*(2), 103-113.

Martinez, R. G., Lewis, C. C., & Weiner, B. J. (2014). Instrumentation issues in implementation science. *Implementation Science, 9*, 1-9.

McKay, M. M., & Bannon, W. M. (2004). Engaging families in child mental health services. *Child and Adolescent Psychiatric Clinics of North America*, *13*, 905-921.

McLeod, B. D. (2011). Relation of the alliance with outcomes in youth psychotherapy: A meta-analysis. *Clinical Psychology Review, 31*, 603-616.

McLeod, B. D., Islam, N. Y., Chiu, A. W., Smith, M. M., Chu, B. C., & Wood, J. J. (2014). The relationship between alliance and client involvement in CBT for child anxiety disorders. *Journal of Clinical Child & Adolescent Psychology, 43*(5), 735-741.

McLeod, B. D., Islam, N. Y., & Wheat, E. (2013). Designing, conducting, and evaluating therapy process research. In J. Comer & P. Kendall (Eds.), *The Oxford Handbook of Research Strategies for Clinical Psychology* (pp. 142-164). New York, NY, US: Oxford University Press.

McLeod, B.D., Smith, M. M., Southam-Gerow, M. A., Weisz, J. R., & Kendall, P. C. (2015). Measuring treatment differentiation for implementation research: The therapy process observational coding system for child psychotherapy revised strategies scale. *Psychological Assessment, 27*(1), 314-325.

McLeod, B.D., Southam-Gerow, M.A., & Kendall, P.C. (2017). Observer, youth, and therapist perspectives on the alliance in cognitive behavioral treatment for youth anxiety. *Psychological Assessment*. Advance online publication.

McLeod, B. D., Southam-Gerow, M. A., Rodriguez, A., Quinoy, A. M., Arnold, C. C., Kendall, P. C., & Weisz, J. R. (2016). Development and initial psychometrics for a therapist competence instrument for CBT for youth anxiety. *Journal of Clinical Child and Adolescent Psychology*. Advanced online publication.

McLeod, B. D., & Weisz, J. R. (2005). The Therapy Process Observational Coding System – Alliance Scale: Measure characteristics and prediction of outcome in usual clinical practice. *Journal of Consulting and Clinical Psychology*, *73*(2), 323-333.

McLeod, B. D., & Weisz, J. R. (2010). The Therapy Process Observational Coding System for Child Psychotherapy Strategies Scale. *Journal of Clinical Child & Adolescent Psychology, 39*, 436-443.

Merikangas, K. R., He, J., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., … Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: Results from the national comorbidity survey replication – adolescent supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry, 4* (10), 980-989.

Morris, E., Fitzpatrick, M. R., & Renaud, J. (2014). A pan-theoretical conceptualization of client involvement in psychotherapy. *Psychotherapy Research*, *26*, 70-84.

Nichols, M. P. (2006). *Family therapy: Concepts and methods* (7th ed.). Boston: Allyn & Bacon.

Nichols, M. P. (2007). *The essentials of family therapy* (3rd ed.). Boston: Allyn & Bacon.

Nix, R. L., Bierman, K. L., & McMahon, R. J. (2009). How attendance and quality of participation affect treatment response to parent management training. *Journal of Consulting and Clinical Psychology*, *77*(3), 429-438.

Noser, K. & Bickman, L. (2000). Quality indicators of children's mental health services: Do they predict improved client outcomes?. *Journal of Emotional and Behavioral Disorders, 8*(1), 9-18.

O'Malley, S. S., Suh, C. S., & Strupp, H.H. (1983). The Vanderbilt psychotherapy process scale: A report on the scale development and a process-outcome study. *Journal of Consulting and Clinical Psychology*, *51*(4), 581-586.

Orlinsky, D. E., Grawe, K., & Parks, B. K. (1994). Process and outcome in psychotherapy: Noch einmal. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed.), (pp.270-376). Oxford, England: John Wiley & Sons.

Richards, M. M., Bowers, M. J., Lazicki, T., Krall, D., & Jacobs, A. K. (2008). Caregiver involvement in the intensive mental health program: Influence on changes in child functioning. *Journal of Child and Family Studies*, *17*, 241-252.

Rogers, C. R. (1987). The underlying theory: Drawn from experiences with individuals and groups. *Counseling and Values*, *32*(1), 38-45.

Rosenthal, R. & Rosnow, R. L. (1984). *Essentials of Behavioral Research: Methods and Data Analysis.* New York: McGraw-Hill.

Shavelson, R.J. Webb, N.M. & Rowley, G.L. (1989). Generalizability theory. *American Psychologist*, *44*(6), 922-932.

Shelef, K. & Diamond, G. M. (2008). Short form of the revised Vanderbilt therapeutic alliance scale: Development, reliability, and validity. *Psychotherapy Research, 14*(4), 433-443.

Shirk, S. R., Crisostomo, P. S., Jungbluth, N., & Gundmundsen, G. R. (2013). Cognitive mechanisms of change in CBT for adolescent depression: Associations among client involvement, cognitive distortions, and treatment outcome. *International Journal of Cognitive Therapy, 6*(4), 311-324.

Shirk, S. R., & McMakin, D. (2008). Client, therapist, and treatment characteristics in EBTs for children and adolescents. In R. G. Steele, T. D. Elkin, & M. C. Roberts (Eds.), *Handbook of evidence-based therapies for children and adolescents: Bridging science and practice. Issues in clinical child psychology* (pp. 471-486). New York, NY: Spring Science + Business Media.

Shirk, S. R., & Saiz, C. C. (1992). Clinical, empirical, and developmental perspectives on the therapeutic relationship in child psychotherapy. *Development and Psychopathology, 4*, 713-728.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420–8.

Silverman, W. & Albano, A. M. (1996). *The Anxiety Disorders Interview Schedule for DSM-IV: Child and parent versions.* San Antonio, TX: Graywind.

Southam-Gerow, M. A. & McLeod, B. D. (2013). Advances in applying treatment integrity research for dissemination and implementation science: Introduction to special issue [Special issue]. *Clinical Psychology: Science and Practice, 20*, 1-13.

Southam-Gerow, M. A., Weisz, J., Chu, B. C., McLeod, B., Gordis, E., & Connor-Smith (2010). Does cognitive behavioral therapy for youth anxiety outperform usual care in community

clinics? An initial effectiveness test. *Journal of the American Academy of Child and Adolescent Psychiatry, 49*(10), 1043-52.

Suh, C. S., O'Malley, S. S., Strupp, H. H. & Johnson, M. E. (1989). The Vanderbilt psychotherapy process scale (VPPS). *Journal of Cognitive Psychotherapy: An International Quarterly, 3*, 123-154.

Tichenor, V., & Hill, C. E. (1989). A comparison of six measures of working alliance. *Psychotherapy, 26*, 195-199.

Toban, J. I., Eichstedt, J. A., Wolfe, V. V., Phoenix, E., Brisebois, S., Zayed, R. S., … Harris, Kimberly E. (2011). Group cognitive-behavioral therapy for anxiety in a clinic setting: Does child involvement predict outcome? *Behavior Therapy, 42*, 306-322.

Tolan, P. H., Hanish, L. D., McKay, M. M., & Dickey, M. H. (2002). Evaluating process in child and family interventions: Aggression prevention as an example. *Journal of Family Psychology, 16*(2), 220-236.

Tryon, G. S. & Winograd, G. (2011). Goal consensus and collaboration. *Psychotherapy, 48*(1), 50-57.

Weisz, J. R., Chorpita, B. F., Palinkas, L. A., Schoenwald, S. K., Miranda, J., Bearman, S. K., …the Research Network on Youth Mental Health. (2012). Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth: A randomized effectiveness trial. *Archive of General Psychiatry, 69*(3), 274-282.

Weisz, J. R., Moore, P. S., Southam-Gerow, M. A., Weersing, V. R., Valerie, S. M., & McCarty, C. A. (2005). *Therapist's manual PASCET: Primary and Secondary Control Enhancement Training Program* (3rd ed.). Los Angeles: University of California.

Weisz, J. R., Southam-Gerow, M. A., Gordis, E. B., Connor-Smith, J. K., Chu, B. C., Langer, D. A., … Weiss, B. (2009). Cognitive-behavioral therapy versus usual clinical care for youth depression: An initial test of transportability to community clinics and clinicians. *Journal of Consulting and Clinical Psychology*, *77*(3), 383-396.

Weller, E. B., Weller, R. A., Rooney, M. T., & Fristad, M. A. (1999a). *Children's Interview for Psychiatric Syndromes (ChIPS)*. Washington, DC: American Psychiatric Press.

Weller, E. B., Weller, R. A., Rooney, M. T., & Fristad, M. A. (1999b). *Children's Interview for Psychiatric Syndromes – Parent Version*. Washington, DC: American Psychiatric Association.

Windholz, M. J., Weiss, D. S., & Horowitz, M. J. (1985). An empirical study of the natural history of time-limited psychotherapy for stress response syndromes. *Psychotherapy*, *22*(3), 547-554.

Yalon, I. D. (2003). *The gift of therapy: An open letter to a new generation of therapists and their patients*. New York: HarperCollins (Perennial).

Vita

Emily Jane Butt Wheat was born on October 20, 1983, in Durham, North Carolina. She graduated from Tuscaloosa County High School in Tuscaloosa, Alabama in 2002. She received her Bachelor of Arts in Psychology from the University of South Alabama, Mobile, Alabama in 2006 and her Master of Arts in Psychology from the University of Richmond, Richmond, Virginia in 2009.