VCU

Virginia Commonwealth University

VCU Scholars Compass

Theses and Dissertations

Graduate School

2017

Sticking to the recipe: How do adherence and differentiation to a CBT protocol affect client outcomes in youths with anxiety?

Ruben G. Martinez
Virginia Commonwealth University

Follow this and additional works at: https://scholarscompass.vcu.edu/etd

Part of the Child Psychology Commons, and the Clinical Psychology Commons

Downloaded from

https://scholarscompass.vcu.edu/etd/4728

Sticking to the recipe: How do adherence and differentiation to a CBT protocol affect client outcomes in youths with anxiety?

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science

at Virginia Commonwealth University.

by

Ruben Gabriel Martinez

Bachelor of Arts, Indiana University, 2011

Director, Bryce D. McLeod, Ph.D.

Associate Professor, Department of Psychology

Virginia Commonwealth University

Richmond, VA

April, 2017

Acknowledgement

I would like to thank Dr. McLeod for his endless wisdom, support, guidance, and desire to help me craft this document. In addition, I would like to thank Dr. Cara Lewis, who selflessly provided me with the guidance and mentorship that made admission to graduate school possible. I would also like to thank Lourdes and Romeo Martinez, my incredible parents, for the constant and unending love, support, and dedication. Your sacrifices have allowed for me to follow my dreams.

Table of Contents

Chapter 4

Chapter 5

List of Tables

List of Figures

Abstract

STICKING TO THE RECIPE: HOW DO ADHERENCE AND DIFFERENTIATION TO A CBT PROTOCOL AFFECT CLIENT OUTCOMES IN YOUTHS WITH ANXIETY?

By Ruben G. Martinez, B.A.

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

Virginia Commonwealth University, 2017

Major Director: Bryce D. McLeod,
Associate Professor, Psychology Department

Objective: Understanding the pathways through which treatments work to change symptom and diagnostic outcomes is important to the development and delivery of evidence-based treatments. This study assessed the extent to which adherence (therapist's delivery of prescribed therapeutic interventions) and differentiation (therapist's delivery of non-prescribed therapeutic interventions) to Coping Cat, a CBT program, affected client symptom and diagnostic outcomes. **Method:** The Therapy Process Observational Coding System for Child Psychotherapy – Revised Strategies Scale (McLeod et al., 2015) was used to characterize therapeutic interventions delivered within and outside of the Coping Cat program with youths aged 7-15 receiving treatment in one efficacy ($n = 51$; 41% female; 84% Caucasian, M age = 10.37) and one effectiveness ($n = 17$; 56% female, 39% Caucasian, M age = 10.90) trial. Youth- and parent-report symptom checklists and diagnostic interviews were used to assess symptom and diagnostic remission. Multiple hierarchical regression analyses and hierarchical binomial logistic regression were used to investigate the relation between adherence and differentiation and symptom change and remission of principal diagnosis. **Results:** Neither adherence nor differentiation were significantly related to symptom or diagnostic outcomes. No clear trend emerged, and results were inconsistent across parent and youth report, outcome type, and setting.

Conclusion: These results are consistent with past literature. Two interpretations exist: (1) that there is no relation between treatment delivery and outcomes, and (2) that methodological and analytic flaws undercut the ability of the analyses to identify a relation.

Chapter one

Introduction

Efforts to improve quality of mental healthcare for youth and their families have focused,

in part, on the development and use of evidence-based treatments (EBTs) in both research and

practice settings (American Psychological Association Task Force on Evidence-Based Practice

for Children and Adolescents, 2008; Institute of Medicine, 2015). Many EBTs have been created

for various emotional and behavioral problem areas (Chorpita et al., 2011; Weisz & Gray, 2008).

While the main focus of research on EBTs asks the question of whether or not an EBT works to

change client outcomes (Chambless & Ollendick, 2001), very little research seeks to explain how

treatment works (Weersing & Weisz, 2002). Despite the existence of models that attempt to

explain how treatments work, very few empirical tests of these models exist (Kazdin, 2007). It is

important to understand how treatments work to further the field's understanding of that

treatment and to optimize treatment delivery (Perepletchikova, Treat, & Kazdin, 2007).

One way to study whether therapeutic interventions affect outcomes is through treatment

delivery research, which provides a lens through which researchers can understand how EBTs

work. Treatment delivery refers to a broad domain of research that seeks to explain how

activities that occur in treatment affect client outcomes (McLeod, Islam, & Wheat, 2013). Doss

(2004) created a model that seeks to explain through what pathways treatments work to affect

treatment outcomes. This model theorizes, among other things, how the delivery of therapeutic

interventions should work to affect client treatment outcomes.

Treatment delivery consists of what therapeutic interventions are delivered by the

therapist (i.e., adherence; McLeod, Southam-Gerow, Tully, Rodriguez, & Smith, 2013), how

much the therapist deviates from a treatment model and delivers non-prescribed therapeutic

interventions (i.e., differentiation), and with how much quality interventions are delivered (i.e., competence). Thus, treatment delivery research allows researchers to measure how and what therapeutic interventions are being delivered as well as evaluate relations between therapeutic interventions and client outcomes.

There is a dearth of empirical literature that seeks to demonstrate a relation between treatment delivery and client outcomes. Webb, DeRubeis, and Barber (2010) conducted a meta-analysis that investigated the extent to which adherence and competence affect client outcomes. The review found that the empirical literature up to 2009 showed an inconsistent relation between therapeutic interventions and client outcomes, such that some studies found a strong positive relation, some found a strong negative relation, and others found no relation at all.

One potential hypothesis is that there is no relation between what therapeutic interventions are delivered and client outcomes. However, it is also possible that inconsistent findings were a result of methodological limitations and conceptual inconsistencies; namely, that the characterization and measurement of treatment integrity were inconsistent across reviewed studies and that roughly half of the studies in the Webb et al. (2010) study failed to explore the effects of alternative explanations such as the alliance. Webb et al. (2010) speculate that these, among other methodological, analytic, and conceptual issues directly contributed to inconsistent findings. Conducting a study that addresses the limitations of the extant literature will paint a more comprehensive and convincing picture of the nature and direction of the relation between therapeutic interventions and client outcomes.

The aim of the current study was to understand how the delivery of therapeutic interventions (i.e., adherence and differentiation) affects client outcomes. Though competence is an important aspect of treatment delivery, it would better be investigated as part of a separate

study; thus, competence was excluded from the current study. In particular, this investigation focused upon the extent to which therapeutic interventions prescribed by a cognitive-behavioral therapy (CBT) protocol known as Coping Cat and those not prescribed by Coping Cat affect client outcomes. Client outcome data from two studies, a research-context efficacy study (Kendall et al., 2008), and a practice-context effectiveness study (Southam-Gerow et al., 2010), were used to investigate the effects of the therapeutic interventions on client outcomes (anxiety and remission of anxiety diagnoses) in youths with anxiety. The Therapy Process Observational Coding System for Child Psychotherapy – Revised Strategies (TPOCS-RS; McLeod, Smith, Southam-Gerow, Weisz, & Kendall, 2015) was used to observationally code adherence and differentiation from session recordings of the aforementioned efficacy and effectiveness trials.

It was hypothesized that the extent to which prescribed therapeutic interventions were delivered would correspond with changes in anxiety symptomology and diagnostic status in both settings at post-treatment, and that that relation would remain over and above the effects of non-prescribed interventions and alternative explanations (i.e., client and/or therapist variables; e.g., the alliance). Hierarchical multiple regression and hierarchical binomial logistic regression analyses were used to investigate the relation between the therapeutic interventions deployed, as measured by subscales of the TPOCS-RS, and the youths' symptom and diagnostic remission outcomes.

Chapter two

Literature Review

Youth Mental Health Problems

An estimated 20% of youths in the United States suffer from some form of mental health problem (Soni, 2009). These problems are costly and pervasive (U.S. Department of Health & Human Services, 2011) and can lead to negative short- and long-term effects on physical and mental well-being (Patel, Flisher, Hetrick, & McGorry, 2007). Evidence suggests that a large proportion of youth who receive care for mental health problems do not receive psychological treatments that are supported by evidence (Bickman & Hamilton, 2008). To address this gap, health- and mental health- focused institutes (e.g., Institute of Medicine, 2015; IOM) have made substantial efforts to identify and encourage the use of evidence-based treatments (hereafter referred to as EBTs), which are psychological treatments that have been studied with some level of scientific rigor and have demonstrated positive effects (Hoagwood, Burns, Kiser, Ringelsen, & Schoenwald, 2001).

EBTs for Youth Mental Health

Over the last three decades, the American Psychological Association and IOM have made clear the need for identification and use of EBTs that work to resolve youth mental health problems (American Psychological Association Task Force on Evidence-Based Practice for Children and Adolescents, 2008; IOM, 2015). Much progress has been made, and many EBTs exist for a variety of mental health problems with varying levels of evidence (Chorpita et al., 2011; Weisz & Gray, 2008) as to their efficacy (i.e., an EBT demonstrates positive effects in an efficacy trial; Weisz, Ugueto, Cheron, & Herren, 2013). EBTs are usually developed and initially evaluated in highly controlled research (e.g., university) settings (Schoenwald & Hoagwood, 2001). Data show that some EBTs are efficacious for treating youth mental health problems (e.g.,

Chorpita et al., 2011; Eyberg, Nelson, & Boggs, 2008; Pelham & Fabiano, 2008; Silverman & Hinshaw, 2008; Silverman, Pina, & Viswesvaran, 2008).

A meta-analysis by McLeod and Weisz (2004) calculated effect sizes for efficacy studies that were part of published studies and dissertations. The authors presented *d*-type effect size estimates for 121 dissertations and 134 published studies. Following Cohen (2013), an effect size of .20 is small, .50 is medium, and .80 is large. The mean effect sizes of the published studies and dissertation studies were .50 and .23, indicating a medium and small effect, respectively. These results indicate that EBTs are, at the least, efficacious.

How are EBTs studied?

Extensive work has been done to test the efficacy of EBTs (Chorpita et al., 2011). This is often done by using methodology that examines the effect of the entire package of interventions congruent with that EBT on outcomess, meaning that studies often ask the question of "does this EBT work to change client outcome X," as opposed to the question of "through what pathways does this EBT work to change client outcome X?" (Kazdin, 2001, 2007; Weersing & Weisz, 2002). Many theories and models have been created that seek to explain the pathways through which EBTs work (e.g., Doss, 2004; McLeod, Islam, et al., 2013). Despite the work to create those models, very few have been tested empirically (Weersing & Weisz, 2002). It is necessary to test these models, because while EBTs for youth mental health problems are efficacious, we must understand through what pathways (i.e., mechanisms or processes) EBTs works in order to further our understanding and refinement of the creation and delivery of EBTs (Kazdin, 2007). One predominant method of attempting to understand these pathways is through treatment process research.

**The Treatment Process, Treatment Delivery, and EBTs**

The treatment process is multi-faceted and can be defined as the activities and behaviors that occur in treatment sessions and how these things influence client outcomes (McLeod, Islam, et al., 2013). The treatment process can be studied through assessing treatment delivery. One way to characterize treatment delivery, or the components of treatment that have to do with the therapist (Schulte et al., 2009), involves treatment integrity. Treatment integrity is comprised of what therapeutic interventions prescribed by the treatment were delivered (i.e., adherence; McLeod, Southam-Gerow, et al., 2013), whether delivered interventions are not prescribed by a specific EBT (i.e., differentiation; McLeod, Southam-Gerow, et al., 2013), and quality of treatment delivery, or how skillfully the prescribed interventions were delivered (i.e., competence; McLeod, Southam-Gerow, et al., 2013).

A number of conceptual models exist that seek to characterize how the delivery of therapeutic interventions (i.e., adherence and differentiation) relates to client outcomes (e.g., Doss, 2004; McLeod, Islam, et al., 2013). The Doss (2004) model postulates that the delivery of prescribed therapeutic interventions works to change client outcomes in the targeted domain through pathways of change mechanisms and processes. Figure 1 depicts the Doss (2004) treatment process model. The large blue arrow represents the therapeutic interventions delivered by the clinician. For instance, CBT for youth anxiety (an EBT) contains a set of prescribed CBT-specific therapeutic interventions. The Doss model hypothesizes that the delivery of CBT-specific therapeutic interventions (e.g., psychoeducation on cognitive restructuring) disrupts the pathological process by altering change mechanisms (i.e., changes in client characteristics or skills that lead to improvements; e.g., threat interpretation) through repeated use and refinement of client change processes (i.e., client behaviors or experiences that come about as a result of

delivered therapeutic interventions; e.g., use of cognitive restructuring skills in session), leading to improved client symptom or diagnostic outcomes in the domain of anxiety. It is important to understand whether the delivery of CBT interventions decreases client symptomology through the intended pathways.

The logic of the Doss (2004) model also extends to the delivery of non-prescribed therapeutic interventions. Non-prescribed therapeutic interventions may affect some targeted outcome(s) through change mechanisms, but if the intervention does not target change mechanisms that are believed to affect the targeted outcome, there should be little to no effect of the non-prescribed therapeutic intervention on that outcome. Thus, it would be expected that delivery of therapeutic interventions specifically targeting youth anxiety (i.e., prescribed therapeutic interventions, in this case) would have more of an effect on client anxiety outcomes than those that do not target youth anxiety, unless those interventions that do not target youth anxiety target the same change mechanism. It is important to understand whether this is the case, as little research extends treatment delivery research to include non-prescribed interventions, which could work to augment or dilute the effects of treatment (Perepletchikova & Kazdin, 2005).

FIGURE 1
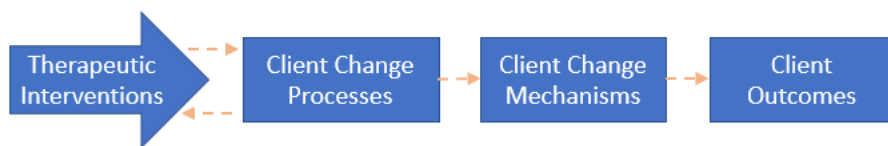
The Doss (2004) Treatment Process Model

The Doss (2004) model makes sense intuitively; however, the relative lack of model testing in psychological treatment research necessitates empirical study of the model to refine

our understanding of the way that treatment, and treatment delivery in particular, works to change client outcomes (Kazdin, 2007; Weersing & Weisz, 2002). To ascertain whether the aforementioned relation (i.e., [A] delivery of prescribed interventions [B] alters change mechanisms to [C] decrease symptoms) exists, it is important first to know that there is a relation between delivered therapeutic interventions (i.e., treatment delivery) (A) and client outcomes (C) (Doss, 2004). To test this model, it was necessary to measure treatment delivery of prescribed and non-prescribed therapeutic interventions (adherence and differentiation, respectively), and assess how the delivery of those therapeutic interventions affected youth symptom and diagnostic outcomes. While few studies exist that speak to the relation between treatment differentiation and client outcomes, a number of studies have sought to answer the question of whether treatment adherence affects client outcomes.

Delivery of Therapeutic Interventions and Client Outcomes

A number of studies exist that explore the relation between treatment delivery and client outcomes, mainly through assessing adherence to prescribed therapeutic interventions. It is important to note that while assessing for the quality of what was delivered (i.e., competence) is important, it is a separate research question and thus was not addressed in this review or study. Webb et al. (2010) performed a meta-analysis that synthesized the literature focused on the relation between therapist adherence and competence and client outcomes. Thirty-two identified studies examined the relation between therapist adherence and competence and client outcomes conducted up until 2009. This review focused primarily on studies that sought to investigate a link between adherence and client outcomes.

Webb et al. (2010) defined adherence as the degree to which therapists delivered therapeutic interventions specifically related to the theory that drives an intervention. Webb and

colleagues calculated *r*-type effect size estimates based upon the methods of Rosenthal, Hoyt,

Ferrin, Miller, and Cohen (2006); these effect sizes speak to the extent to which adherence was

related to client outcomes. For this type of estimate, an effect size of .10 is considered small, .24

is considered medium, and .37 is considered large (Rosnow & Rosenthal, 1996). The effect sizes

for studies examining this relation ranged from -.40 to .47 with a mean effect size of .02, which

was not significantly different from 0. These results indicate that some studies found a large

relation between treatment adherence and outcome where symptoms get better, some found a

large relation where symptoms got worse, and others found no relation at all.

Webb and colleagues speculate that there are two possible reasons for the mixed findings.

The first hypothesis is that there is no relation between adherence to prescribed therapeutic

interventions and client outcomes. The second hypothesis posits that there are definitional and

methodological gaps (e.g., not controlling for confounding variables such as the alliance) in the

majority of the analyzed studies which may manifest in uninterpretable or incomparable effect

sizes. Taken together, the literature that Webb et al. reviewed regarding the relation between

delivered therapeutic interventions and client outcomes is inconsistent and confounded by a

number of important factors.

An updated review

Since the time of the Webb et al. (2010) review, few studies have been conducted that

seek to determine the relation between treatment delivery and client outcomes. In order to

identify meaningful themes and changes to the literature since the Webb et al. review, a literature

review was conducted to identify studies conducted between 2010 to 2015 by using the same

search terms as Webb and colleagues. Search terms were entered into PsycINFO as follows:

therapist OR therapy OR psychotherapist OR psychotherapy AND adherence OR adhering OR

adherent OR adhere OR competence OR competency OR competencies OR integrity. These search terms were expanded by including the term "differentiation." The search included English-language empirical studies only. Following the Webb et al. (2010) review, studies were included that: (a) quantitatively assessed treatment delivery and outcome variables, (b) delivered individual treatment as opposed to group or family treatment, (c) used audiotaped, videotaped, or transcribed treatment sessions rated by trained experts as opposed to therapist or client raters, (d) included a clinical population, and (e) had a sample size of at least five clients. This review focused primarily on studies that assessed the relation between adherence and differentiation and client outcomes because the current study was focused upon treatment delivery (i.e., adherence and differentiation), as opposed to quality of treatment delivery (i.e., competence). The purpose of this review was to identify key methodological and conceptual themes.

After excluding studies based upon the criteria listed above, five studies remained for review (Becker, Becker, & Ginsburg, 2012; Boswell et al., 2013; Liber et al., 2010; Owen & Hilsenroth, 2014; Webb et al., 2012). Of the five studies, three focused upon the effect of treatment delivery on anxiety outcomes (Becker et al., 2013; Boswell et al., 2013; Liber et al., 2010), one on depressive symptoms (Webb et al., 2012), and one on a broad array of client-reported symptomology (e.g., somatic complaints, depression symptoms, anxiety symptoms; Owen & Hilsenroth, 2014). Two studies focused upon youths (Becker et al., 2013; Liber et al., 2010) and three studies focused upon adults (Boswell et al., 2013; Owen & Hilsenroth, 2014; Webb et al., 2012).

Treatment delivery definition. In the reviewed literature, treatment delivery was defined primarily as adherence to prescribed interventions, which excludes the collection and analysis of data that speak to the effects of non-prescribed interventions). While there has been

10

some speculation on the extent to which non-prescribed therapeutic interventions affect client outcomes (e.g., McLeod, Southam-Gerow, et al., 2013), there is a lack of empirical literature that speaks to the effects, harmful or beneficial, of treatment differentiation (i.e., use of non-prescribed interventions; Perepletchikova et al., 2007; Perepletchikova, 2009). It is worth noting that few identified studies sought to assess the effects of treatment differentiation, and those with that aim did not meet inclusion criteria for this review. By limiting the scope of treatment delivery to adherence to prescribed interventions, as opposed to identifying and studying what is delivered within and outside the EBT, the field may be painting an incomplete picture of how treatment delivery affects outcomes (McLeod et al., 2015). In addition, the lack of consistent definition of treatment delivery, and in this case adherence, makes cross-study comparison difficult.

All studies in the current review defined adherence as the extent to which a therapist delivers interventions prescribed by the EBT protocol or manual (Becker et al., 2013; Boswell et al., 2013; Liber et al., 2010; Owen & Hilsenroth, 2014; Webb et al., 2012). No studies that met criteria for inclusion took into account differentiation or directly assessed the effects of differentiation on outcomes. Because some literature suggests that clinicians may deliver interventions outside the scope of a given protocol (Weisz et al., 2009), and because there is little to no literature that assesses the effects of non-prescribed interventions, the field may not be accounting for interventions being delivered outside an EBT or the possibility that non-prescribed interventions may augment or dilute the effects of a treatment (Perepletchikova & Kazdin, 2005). To better understand how the delivery of therapeutic interventions affects outcome, and further to understand how to refine the delivery of EBTs, it is necessary to

11

understand the effects of both prescribed and non-prescribed therapeutic interventions by including these in measurement efforts and analyzing their effects (Perepletchikova et al., 2007).

Treatment delivery measurement. There are many ways to measure the delivery of therapeutic interventions. To understand the relation between treatment delivery and its predicted outcome, it is important to use a measurement strategy that is a comprehensive representation of the interventions that were delivered (McLeod, Islam, et al., 2013). If the strategy used to measure treatment delivery is not representative of what therapeutic interventions are delivered, or to what extent those interventions are delivered, there is a chance that important aspects of treatment delivery are not being accounted for in the obtained results (McLeod, Islam, et al., 2013).

The way that treatment delivery is defined in the reviewed studies directly speaks to the way that it is characterized and measured. The simplest way to measure treatment delivery is by measuring whether or not the components of an EBT are present or absent (e.g., Becker et al., 2013; Boswell et al., 2013). This method assesses whether the therapeutic interventions prescribed by an EBT are or are not delivered to the client. Another way to measure treatment delivery is by assessing how similar the delivered therapeutic interventions were to therapeutic interventions expected to be delivered as part of a given treatment (e.g., Owen & Hilsenroth, 2014). The most complex method of measuring treatment delivery is through an extensiveness score (e.g., Liber et al., 2010; Webb et al., 2012), which combines frequency and thoroughness and speaks to the extent to the depth of a delivered intervention (Hogue et al., 1996). All of the aforementioned measurement methods were represented in the literature, so it is clear that treatment delivery measurement methods are being used inconsistently across studies.

The use of any one measurement strategy is not inherently good or bad. However, the inconsistent use of these strategies indicates that different questions may be being answered across each measurement strategy (Waltz, Addis, Koerner, & Jacobson, 1993). For instance, a presence/absence count does not provide a comprehensive picture of a treatment session, as it does not take into account the frequency and thoroughness (i.e., extensiveness) with which therapeutic interventions are delivered. The presence/absence measurement asks the question, "does the presence or absence of prescribed therapeutic interventions affect client outcomes?", while the extensiveness measurement can be equated to how strong of a dose of the intervention was given. Simple frequency counts increase the possibility that an intervention that was delivered numerous times, but not thoroughly will correspond to change when that is not necessarily the case (McLeod, Islam, et al., 2013). In other words, a CBT-prescribed intervention such as cognitive restructuring may be mentioned frequently within a treatment session, but if it is not thoroughly explained to or explored with the client, the client may not understand the rationale or guiding principles and therefore not engage in cognitive restructuring activities or action plans.

The inconsistency with which treatment delivery is measured creates problems when trying to compare across studies and create a cohesive literature base, and could be a reason why mixed findings were observed in the Webb et al. (2010) meta-analysis. To resolve these issues within a correlational framework, adherence and differentiation should be measured with a scoring strategy that takes into account both breadth and depth of what is being delivered, as this provides the most accurate depiction of the extent to which prescribed and non-prescribed interventions are delivered (McLeod, Islam, et al., 2013), addressing the limitations of the presence/absence and frequency count methods. Frequency and presence/absence scoring

13

strategies are not comprehensive, as they do not take into account thoroughness with which interventions are delivered. Thus, in this case, the use of an extensiveness scoring system that takes into account presence and absence, frequency, and thoroughness of delivered interventions was the most appropriate (McLeod, Islam, et al., 2013).

Treatment delivery, alternative explanations, and client outcomes. Particularly in the case of correlational studies where establishing causality is not possible, it is important to assess for the effects of alternative (i.e., third variable) explanations. Observational studies that seek to explain the relation between treatment delivery and client outcomes have historically done a poor job of assessing and controlling for alternative explanations (Feeley, DeRubeis, & Gelfand, 1999; Webb et al., 2010). Alternative explanations refer to variables that are not measured that may confound or otherwise explain the relation between the independent variable and dependent variable (Webb et al., 2010). Various alternative explanations have been postulated and studied (e.g., client characteristics; McLeod, Southam-Gerow, et al., 2013) that may explain the relation between treatment delivery and client outcomes.

Webb et al. (2010) and others (Barber et al., 2006; McLeod, Southam-Gerow, et al., 2013) have identified the alliance (defined as "the collaborative and affective bond between therapist and patient"; Martin, Garske, & Davis, 2000) as an important third variable that can affect the relation between treatment delivery and client outcomes. Not assessing and controlling for the alliance can increase the possibility of finding a relation between treatment delivery and outcome when there is no relation (Webb et al., 2010). A number of the reviewed studies assessed or controlled for the alliance (Liber et al., 2010; Owen & Hilsenroth, 2014; Webb et al., 2012), while others did not (Becker et al., 2013, Boswell et al., 2013). More consistent

measurement of alliance is a step forward in the field, and is necessary in establishing a relation between treatment delivery and outcome.

No reviewed studies (e.g., Becker et al., 2013; Boswell et al., 2013; Liber et al., 2010; Owen & Hilsenroth, 2014; Webb et al., 2012) included possible alternative explanations beyond the alliance that could be responsible for observed relations between treatment delivery and outcomes (e.g., patient characteristics; Webb et al., 2010). Particularly in correlational studies, it is important to rule out the possibility that third variables such as therapist experience are responsible for finding a relation when there is no relation (Perepletchikova & Kazdin, 2005). To do this, studies should assess for variables such as these and identify their effect on the relation between treatment delivery and client outcomes by including them in statistical analyses (Webb et al., 2010). Consistent measurement and consideration of third variable explanations is crucial for research that seeks to identify a relation between treatment delivery and outcomes.

Establishing temporal precedence. Establishing temporal precedence is an important consideration when attempting to establish a relation between treatment delivery and outcomes (Feeley et al., 1999; Kazdin, 2008). Establishing temporal precedence means that the predicting, or independent variable (treatment delivery, in this case), should be measured temporally before the outcome variable (client outcomes, in this case) that it is meant to predict (Feeley et al., 1999; Judd & Kenny, 1981). If the requirements for temporal precedence are not met, it is difficult to say with confidence that (a) the predicting variable is related to the dependent variable and (b) third variables do not explain the relation between intervention(s) X and client outcome Y (Judd & Kenny, 1981).

All of the studies reviewed for the current study established temporal precedence (Becker et al., 2013; Boswell et al., 2013; Liber et al., 2010; Owen & Hilsenroth, 2014; Webb et al.,

15

2012). Thus, it seems that establishing temporal precedence has emerged as a norm for research that seeks to identify a relation between treatment delivery and outcomes. The only way to establish temporal precedence is by ensuring that the measurement of the independent variable precedes the measurement of the dependent variable in time (Feeley et al., 1999). Conducting a study that meets these requirements will help to rule out the possibility that changes in the client outcomes are a result of temporal sequencing of the measurement of variables.

Updated Summary of the Literature. Methodological, conceptual, and definitional factors play a role in the confidence that can be placed in the reviewed studies. First, treatment delivery was typically narrowly defined as adherence to prescribed therapeutic interventions, which excluded the collection and analysis of data that speak to non-prescribed interventions. Second, treatment delivery was often characterized or measured inconsistently across studies. Third, many studies did not attempt to identify the effects of third variables such as the alliance. Finally, establishing temporal precedence was identified as an emerging norm in this area of research. All of these factors may play a role in the extent to which findings from the reviewed studies are generalizable to the literature at large. Some reviewed articles found a relation between treatment delivery and outcomes (Owen & Hilsenroth, 2014), while others found no relation (Becker et al., 2013; Boswell et al., 2013, Liber et al., 2010) or a mixed relation (Webb et al., 2012), indicating that the findings in this literature are still inconclusive. The growing literature base that attempts to answer the question of whether, and to what extent, treatment delivery affects outcomes is inconclusive.

The Present Study

The present study aimed to identify a relation between treatment delivery, and in particular adherence and differentiation, and client outcomes by addressing the gaps in the

reviewed literature. The treatment that was focused on in this study is Coping Cat (Kendall & Hedtke, 2006 a,b), a 16-20 session, manualized, and individualized CBT (ICBT) protocol used to reduce anxiety symptoms and ultimately diagnoses of youth anxiety. In particular, data from one study that investigated the efficacy of Coping Cat in a research setting (Kendall et al., 2008) was used. Coping Cat has additionally been studied in practice setting; the Youth Anxiety Study (YAS) assessed the effectiveness of Coping Cat in community-based mental health centers (Southam-Gerow et al., 2010). The use of data from these two studies allowed for the exploration of the relation between treatment delivery and client outcomes in the same EBT in both a research and practice setting. If the rationale put forward in the Doss (2004) model is correct, and outcomes are changed through pathways spurred along by the delivery of therapeutic interventions, then this relation should remain consistent across settings. This was an opportunity to study the extent to which this is the case.

A great deal of the literature to date has only analyzed the relation between prescribed therapeutic interventions (i.e., adherence) and client outcomes. Due to the paucity of research in this area, as well as the rationale put forth in the literature review, this research extended on past research by assessing the relation between both prescribed and non-prescribed therapeutic interventions and outcomes. As recommended in McLeod, Islam, et al. (2013) and done in Southam-Gerow et al. (2010), an instrument was used that is capable of broadly assessing the extensiveness with which interventions within and outside the scope of the given protocol were delivered. The Therapy Process Observational Coding System for Child Psychotherapy Revised Strategies Scale (TPOCS-RS; McLeod et al., 2015) is an observational coding instrument capable of assessing treatment differentiation (McLeod, Southam-Gerow, et al., 2013). The TPOCS-RS contains 46 items grouped into five theory-based subscales (cognitive, behavioral,

psychodynamic, family, and client-centered). This instrument does not limit the definition of treatment delivery to adherence to one EBT; however, because items on the TPOCS-RS encompass those expected to be delivered in CBT, the TPOCS-RS is a good candidate for assessing adherence to CBT. The TPOCS-RS allowed for the characterization of what prescribed and non-prescribed treatments were used and to what extent the interventions affected client outcomes. Adherence to prescribed therapeutic interventions was assessed with methods from Southam-Gerow et al. (2010), who created a subscale of the TPOCS-RS that corresponds to the prescribed therapeutic interventions of Coping Cat. To represent differentiation, two subscales (TPOCS-RS Family and Psychodynamic subscales) of the TPOCS-RS were used that contained therapeutic interventions not expected to be delivered in Coping Cat.

Hoagwood, Jensen, Petti, and Burns (1996) presented a model of outcomes that are important to measure in youth mental health treatment. Two outcomes of particular interest were symptoms and diagnosis, both intended targets of treatment. According to Hoagwood et al., symptoms are defined as emotional or behavioral manifestations demonstrated in one or more settings, while diagnosis refers to the patterns of that symptomology and whether that pattern is consistent with criteria presented in a diagnostic manual (e.g., Diagnostic and Statistical Manual of Mental Disorders or the International Statistical Classification of Diseases and Related Health Problems). In particular, this study focused upon: (a) the extent to which delivered therapeutic interventions reflected decreases in youth anxiety symptoms that Coping Cat was designed to target and (b) the extent to which delivered therapeutic interventions corresponded to diagnostic remission of a principal anxiety diagnosis. The temporal ordering of these data (i.e., treatment delivery was measured before outcome) allowed for the establishment of temporal precedence, as indicated in Feeley and DeRubeis (1999).

Assessing the relation between therapeutic interventions and their targeted outcome(s) was a difficult task due to methodological and definitional inconsistencies in the literature. In addition, there was a dearth of empirical literature that directly spoke to the strength and direction of this relation. This study focused upon treatment integrity data from an observational coding system and pre- and post-treatment outcome data from Coping Cat delivered in research and community settings to assess to what extent treatment delivery was related to client symptom outcomes in youths with anxiety disorders.

Hypotheses

It was hypothesized that delivered therapeutic interventions would be related to anxiety symptom outcomes (i.e., continuous outcome variables) in both studies. In particular, it was hypothesized that adherence (i.e., delivery of prescribed therapeutic interventions measured by TPOCS-RS Coping Cat subscale), would correspond to client symptom outcome, defined as anxiety symptomology (measured by Child Behavior Checklist-Anxiety scale and Multidimensional Anxiety Scale for Children), in the Kendall study (hypothesis 1a). It was also hypothesized that this relation would remain significant over and above the effects of alternative explanations and differentiation (i.e., delivery of non-prescribed therapeutic interventions measured by TPOCS-RS Family and Psychodynamic subscale; hypothesis 1b).

With regard to the YAS trial, it was hypothesized that adherence (i.e., delivery of prescribed therapeutic interventions; measured by TPOCS-RS Coping Cat subscale) would correspond to client symptom outcomes (measured by YAS youth and parent factor scores; hypothesis 2a), and additionally that the relation between adherence and client outcomes would remain significant over and above the effects of alternative explanations and differentiation (i.e.,

delivery of non-prescribed therapeutic interventions measured by TPOCS-RS Family and Psychodynamic subscale; hypothesis 2b).

It was also hypothesized that delivered therapeutic interventions would be related to diagnostic remission in both studies. In particular, it was hypothesized that adherence (i.e., delivery of prescribed therapeutic interventions; measured by TPOCS-RS Coping Cat subscale), would correspond to remission of principal anxiety diagnoses, defined as anxiety symptomology (ADIS-C/P), in the Kendall study (hypothesis 3a). It was also hypothesized that this relation would remain over and above the effects of alternative explanations and differentiation (i.e., delivery of non-prescribed therapeutic interventions measured by TPOCS-RS Family and Psychodynamic subscale; hypothesis 3b).

With regard to the YAS trial, it was hypothesized that adherence (i.e., delivery of prescribed therapeutic interventions measured by TPOCS-RS Coping Cat subscale) would correspond to remission of principal anxiety diagnosis (measured by DISC 4.0; hypothesis 4a), and additionally that the relation between adherence and diagnostic remission would remain significant over and above the effects of alternative explanations and differentiation (i.e., delivery of non-prescribed therapeutic interventions measured by TPOCS-RS Family and Psychodynamic subscale; hypothesis 4b).

Chapter 3

Method

Participants

This study used data from two randomized clinical trials that provided treatment to youth with anxiety disorders. The first trial, conducted by Kendall et al. (2008; hereafter referred to as Kendall), was conducted at Temple University (Kendall, Hudson, Gosch, Flannery-Schroder, & Suveg, 2008). Kendall et al. (2008) compared the efficacy of family CBT, ICBT, and family-based education and support/attention. The second trial, called the Youth Anxiety Study (hereafter referred to as YAS), was conducted in Los Angeles County (Southam-Gerow et al., 2010). The YAS sought to compare the effectiveness (i.e., evidence from a trial that uses clinically representative treatment groups and conditions; Weisz, Jensen-Doss, & Hawley, 2006) of ICBT versus usual care (i.e., treatment, counseling, or case management that would be provided as part of a routine service by mental health centers or other providing organizations; UC; Weisz et al., 2006) in six outpatient community mental health clinics. The current study drew from the ICBT condition of the Kendall et al. (2008) study ($N = 55$) and the ICBT condition of the YAS ($N = 24$). Youths were excluded from this study if they received treatment from multiple therapists or if treatment delivery (i.e., TPOCS-RS) data were available for fewer than two sessions, as two data points (i.e., two sessions) are necessary to analyze these data. For these reasons, $n = 4$ youths were excluded for this reason in the Kendall et al. (2008) study (final $n = 51$) and $n = 7$ participants from the YAS (final $n = 17$). See Table 1 for descriptive demographic data regarding the Kendall et al. (2008) and YAS subsamples.

Table 1

Descriptive data of Kendall et al. (ICBT condition) and YAS samples

Variable	Kendall	YAS
Youth age in years	10.37 *(1.88)*	10.90 *(2.10)*
Percent of female youths	41.80	56.20
Percent of Caucasian youths	83.64	38.50
Percent of youths with principle diagnoses		
GAD	47.00	12.50
SAD	38.00	37.50
SP	-	22.90
SOP	45.00	27.01
Total diagnoses at initial assessment *(SD)*	3.56 *(1.74)*	3.20 *(1.80)*
Therapist age in years *(SD)*	--[a]	33.67 *(9.59)*
Therapist years of training *(SD)*	--[a]	4.40 *(2.20)*
Percent of therapist degrees		
Social Worker	--[a]	27.30
Masters-level		51.50
Doctoral-level		9.10
Other		12.10

Note. ICBT = Individual Cognitive Behavioral Therapy condition of Kendall et al. study, YAS = Youth Anxiety Study, GAD = generalized anxiety disorder, SAD = separation anxiety disorder, SOP = social phobia, SP = specific phobia.
In the Kendall et al. trial a number of youths had more than one principal diagnosis.
[a]These data were missing

Kendall et al. (2008) trial. The following section details the youth and therapist

participants in the Kendall et al. (2008) trial.

Youth participants. 51 youths were included in the current subsample of the ICBT

condition of the Kendall study. 61% identified as male and 86% were Caucasian. Reported

family incomes for the youth's families were such that 35% of youths came from families that

earned up to $60,000 dollars annually. Based upon the ADIS-C/P diagnostic interview, the youth

were diagnosed with the following principal anxiety disorders: 37% with Generalized Anxiety

Disorder (GAD), 29% with Separation Anxiety Disorder (SAD), and 33% with Social Phobia

(SOP). It was possible for a youth to have co-occurring principal anxiety disorders if they met

diagnostic criteria and the clinician-assigned severity ratings were equivalent across anxiety

disorders. Comorbidity was common; youths met criteria for an average of 2.49 (SD = 1.17)

anxiety diagnoses.

      Therapist participants. Youth in the ICBT condition received treatment from 16

master's- and doctoral- level (i.e., Ph.D.) therapists. 88% of therapists were female and 68%

were Caucasian. Data regarding therapist age were not reported. Therapists that were not

licensed were supervised by a doctoral-level psychologist with six to seven years of community

experience, while therapists that were licensed were not supervised. Data regarding which

therapists were licensed or not were not reported.

      YAS. The following section details the youth and therapist participants in the YAS.

      Youth participants. A total of 17 youths were included in the current subsample (Coping

Cat; Kendall & Hedtke, 2006 a,b) of the Coping Cat condition of the YAS. Of the 17 participants

that were assigned to ICBT, 71% identified as female and 41% were Caucasian. Reported family

incomes for the youth's families were such that 71% of youths came from families that earned up

to $60,000 annually. Two participants chose not to respond to the demographic survey. Based

upon diagnostic interviews, 6% were diagnosed with GAD, 35% with SAD, and 24% with SOP,

35% for Specific Phobia (SP). Comorbidity was common; a participant met criteria for an

average of 2.00 diagnoses (SD = 1.00) at initial assessment.

      Therapist participants. Youths in the ICBT condition were treated by 13 therapists that

were of social work, master's- and doctoral- level psychology, and other mental health

backgrounds (e.g., marriage and family therapists). 71% identified as female and 41% were Caucasian. Data from two therapists were missing.

 Treatment

Both ICBT conditions used the Coping Cat program (Kendall & Hedtke, 2006 a,b). Coping Cat is a 16-20 session EBT primarily used to treat youth suffering from an anxiety disorder. Coping Cat begins by the therapist providing psychoeducation regarding the youth's anxiety and proceeds into the training phase of treatment (eight sessions), where youths learn about the FEAR acronym. The purpose of the FEAR plan is to teach youths when they are *Feeling* anxious (i.e., identifying an anxious state), when they are *Expecting* negative things to happen (i.e., anxious cognitions), *Actions/Attitudes* that they can use to help manage anxiety (e.g., positive self-talk), and how to identify *Results* (i.e., evaluate performance) and *Reward* themselves for implementing helpful strategies. To encourage and enforce the use of these strategies, therapists use interventions such as modeling, role-playing, homework, and contingencies. The second half of treatment is known as the practice phase of treatment and usually lasts eight or more sessions. The practice phase consists of encouraging youths to apply the skills learned in the training phase to systematic exposure to anxiety-provoking situations or stimuli.

Adherence to Coping Cat was measured in a checklist format in both the Kendall et al. (2008) and Southam-Gerow et al. (2010) trials. This checklist was used to calculate a ratio of the amount of delivered, prescribed Coping Cat interventions over the total amount of interventions prescribed by Coping Cat (i.e., number of Coping Cat interventions used/number of total Coping Cat interventions prescribed for delivery). Kendall et al. (2008) found that 91% of intended prescribed Coping Cat interventions were delivered. Southam-Gerow et al. (2010) found that

99% of sessions contained the expected prescribed interventions. Southam-Gerow et al. (2010) also assessed treatment differentiation using a previous version of the TPOCS-RS, called the Therapy Process Observational Coding System for Child Psychotherapy – Strategies Scale (TPOCS-S), by comparing the extent to which UC versus Coping Cat therapists delivered theory-derived groups of therapeutic interventions. The results indicated that UC therapists versus Coping Cat therapists delivered significantly more psychodynamic, family, and client-centered interventions. This study also used a subscale designed to assess therapeutic interventions found in Coping Cat, called the TPOCS-RS Coping Cat subscale, which includes six interventions specific to Coping Cat. They found that this subscale was able to differentiate between therapists delivering UC and therapists delivering Coping Cat.

Treatment Delivery Instruments

Therapy Process Observational Coding System for Child Psychotherapy – Revised Strategies scale (TPOCS-RS; McLeod et al., 2015) is an observational coding system used to characterize therapeutic interventions delivered by therapists. The TPOCS-RS consists of items that map onto five subscales corresponding to the following theory-based domains: Cognitive (n = 3 items; e.g., *Cognitive Education*), behavioral (n = 8 items; e.g., *Relaxation*), psychodynamic (n = 4 items; e.g., *Interpretation*), family (n = 7 items; e.g., *Parenting Style*), and client-centered (n = 3 items; e.g., *Validates Client*). In addition, there are 13 items not tied to a specific theory-based domain (e.g., *Homework*, *Play Therapy*).

Use of the TPOCS-RS necessitates observation (audio or visual recording) of a treatment session. The TPOCS-RS employs a 7-point likert-type extensiveness rating system (1 = *not at all,* 4 = *considerably*, and 7 = *extensively*) that can be equated to dosage. Extensiveness scores consider (1) how frequently the therapeutic intervention is delivered and (2) with how much

25

thoroughness, or to what depth the therapeutic intervention is delivered (Hogue et al., 1996). For instance, if an intervention like cognitive restructuring (i.e., thought restructuring) was discussed often throughout a session, but the therapist does not address the item in great depth (i.e., does not practice thought restructuring or explain the rationale for its use in great detail), a coder might consider scoring that item a 3 or 4 in extensiveness. On the other hand, if a therapist delivers the item frequently and with substantial depth, a coder may consider scoring the item a 6 or 7. Thus, higher extensiveness scores indicate that therapeutic interventions were delivered thoroughly with high frequency, while lower extensiveness scores indicate that therapeutic interventions were delivered with low frequency and thoroughness.

The TPOCS-RS has evidenced solid psychometric properties. First, the TPOCS-RS has shown evidence of inter-rater reliability across a number of studies, with ICCs ranging from .71-.86 (*M* ICC= .81; McLeod & Weisz, 2010). The TPOCS-RS also has evidenced construct validity in research and practice settings (McLeod & Weisz, 2010; McLeod et al., 2015). Most recently, the TPOCS-RS was able to differentiate between treatment types (McLeod et al., 2015; Southam-Gerow et al., 2010; Weisz et al., 2009; Wood, Piacentini, Southam-Gerow, Chu, & Sigman, 2006). Finally, the TPOCS-RS subscale scores have evidenced predictive validity (Garland et al., 2014). The robust psychometric properties of the TPOCS-RS, in addition to its ability to differentiate between treatments, makes the tool ideal for measuring treatment differentiation.

For the current study, the TPOCS-RS Coping Cat subscale was used to measure prescribed therapeutic interventions (i.e., interventions expected to be delivered as part of Coping Cat). The TPOCS-RS Coping Cat subscale was created by Southam-Gerow et al. (2010) to characterize adherence to therapeutic interventions prescribed by Coping Cat and is made up

of seven items: Cognitive Education, Cognitive Distortion, Coping Skills, Relaxation, Respondent, Operant, and Homework. The average of both coders' scores was taken to create scores on each item and subsequent TPOCS-RS Coping Cat subscale scores.

Two separate TPOCS-RS subscales were used to represent non-prescribed therapeutic interventions, or interventions that are not intended to be delivered as part of ICBT for youth anxiety (Waltz et al., 1993). The first subscale, called the Psychodynamic subscale, consisted of the following five items: Psychodynamic Focus, Transference, Explores Past, Resistance, and Interpretation. The second subscale, called the Family subscale, consisted of the following seven items: General Family Focus, Targets Other Participants, Recruits Others, Parenting Styles, Parenting Skill, Multiparticipant, and Family Roles. The average of both coders' scores was taken to create score on each item and subsequent TPOCS-RS Family and Psychodynamic subscale scores.

Inter-rater reliability was calculated for the relevant TPOCS-RS subscales using Intraclass Correlation Coefficients (ICCs; Shrout, Spitzer, & Fleiss, 1987). ICCs were calculated for the TPOCS-RS Coping Cat, Psychodynamic, and Family subscales. The model ICC (2,2) was the most appropriate for the current study because the same two coders coded each session, and their average ratings served as the unit of analysis. The model dictates reliability coefficients and is based upon a two-way random effects model. Cichetti (1994) specified criteria for *poor* (below .40), *fair* (.40-.59), *good* (.60-.74), and *excellent* (.75 and above) agreement. All TPOCS-RS subscales (Coping Cat, Psychodynamic, and Family) were in the *excellent* range in Kendall and YAS (ICC = .75-.96) with the exception of Psychodynamic subscale in YAS, which was *good* (ICC = .65).

Therapy Process Observational Coding System for Child Psychotherapy – Alliance

scale (TPOCS-A; McLeod & Weisz, 2005) is an observational instrument that assesses the youth- and parent-therapist alliance. The TPOCS-A is comprised of nine items that assess the bond and task dimensions of the alliance. The task items focus on the extent to which clients actively participate in therapeutic activities, while bond items focus on the affective components of the relationship between a parent/youth and the therapist. Direct observation of a treatment session is necessary to code using the TPOCS-A. Coders watched or listened to recordings of treatment sessions and coded behaviors that contributed to the task and bond of the alliance using a designated rating scale (0 = *not at all*; 3 = *somewhat*, 5 = *great deal*). One average TPOCS-A score at the client level was created for use in the present study.

The TPOCS-A has been used in a number of studies (e.g., Chiu, McLeod, Har, & Wood, 2009; Langer, McLeod, & Weisz, 2011). The original study by McLeod and Weisz (2005) showed that scores on the TPOCS-A predicted youth symptom outcomes. The parent-therapist relationship predicted decreases in anxiety symptoms, while the youth-therapist relationship predicted reductions in internalizing, anxiety, and depression symptoms. In another study, Chiu et al. (2009) found that scores on the TPOCS-A, particularly those referring to youth-therapist alliance, predicted client outcomes in youths receiving ICBT at the midpoint of treatment. The TPOCS-A has shown strong psychometric characteristics, including inter-rater reliability (ICC = .48-.80), internal consistency (α = .91-.95), and convergent validity with self-report alliance tools (*r* = .48-.53; Chiu et al., 2009; Fjermestad et al., 2012, McLeod & Weisz, 2005). As such, it was ideal for use in the current study. Cronbach's α for the TPOCS-A in the current study was .84.

Symptomology Instruments

      Child Behavior Checklist (CBCL; Achenbach, 1991) is a 118-item parent-report instrument that assesses a variety of emotional and behavioral symptoms that a youth is experiencing. The CBCL for ages 6-18 (CBCL/6-18) was used for the current study. Parents report the extent to which the youth is experiencing various emotional or behavioral impairments on a three-point scale (0 = *not true;* 1 = *somewhat/sometimes;* 2 = *very/often true)*. A standardized *T* score is generated that is representative of the youth's impairment in a given domain as compared to same-age/sex peers.

      The CBCL consists of three broadband scales (Internalizing, Externalizing, Total) and eight narrowband subscales (Anxious/Depressed, Withdrawn/Depressed, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Rule-Breaking Behavior, and Aggressive Behavior). A *T* score of 65 or over on the broadband scales is considered clinically significant, indicating that a youth may be in need of treatment. Because the intended target of Coping Cat is youth anxiety, the 19-item CBCL Anxiety (CBCL-A) subscale, created by Kendall et al., (2007) was used as the main parent-report outcome of interest. The CBCL-A is highly correlated with other tools meant to measure anxiety in youths (e.g., Multidimensional Anxiety Scale for Children; March, Parker, Sullivan, Stallings, & Conners, 1997), is sensitive to change in youth anxiety, and is able to discriminate between anxious and non-anxious youth (Kendall et al., 2007). Cronbach's alpha for the CBCL-A in the current study was $\alpha = .80$.

      Multidimensional Anxiety Scale for Children (MASC; March et al., 1997) was used in Kendall et al. (2008) as a youth anxiety outcome instrument. The MASC is a 39-item youth self-report scale that assesses anxiety symptoms expected to be targeted by ICBT. The MASC consists of four factors: physical symptoms, social anxiety, harm avoidance, and separation

anxiety. Youths respond as to whether they have been experiencing certain symptoms over the last two weeks on a scale of 1 (*never*) to 4 (*often*). The MASC has demonstrated strong psychometric properties, including test-retest reliability (ICC = .93 at three-month retest), concurrent validity with the ADIS-C/P (Wood, Piacentini, Bergman, McCracken, & Barrios, 2002), and internal consistency, α = .90 (March et al., 1997; March & Sullivan, 1999). For the current study, internal consistency was α = .72.

Anxiety Disorders Interview Schedule for DSM-IV: Child and Parent Versions (ADIS-C/P; Silverman & Albano, 1996) is a semi-structured interview used in Kendall et al. (2008) for assessing for the presence of anxiety disorders. Independent evaluators assessed for the presence of DSM-IV symptom counts and disorders based upon separate interviews with the youth and parents. After these interviews, the evaluator made a severity rating on the ADIS-C/P Clinician's Severity Rating Scale (CSR; 0 = *not at all*, 4 = *some*, 8 = *very, very much*) for each anxiety diagnosis that was assigned. A rating of 4 or greater was considered clinically significant.

Interviews were conducted by doctoral clinical psychology students before, immediately after, and one year following treatment. Evaluators were blind to condition and conducted ADIS-C/P interviews with procedures recommended by ADIS-C/P authors. Evaluators were required to achieve and maintain an inter-rater diagnostic reliability of .85 (Cohen's K; Silverman & Albano, 1996). All sessions were monitored and videotaped to ensure compliance with the ADIS-C/P protocol. The ADIS-C/P has established evidence of test-retest reliability for both the child (κ = .63-.80; ICC = .78-.95) and parent interviews (κ = .65-.92; ICC = .81-.96) across age groups used in this study (Silverman, Saavedra, & Pina, 2001). The ADIS-C/P has also

evidenced concurrent validity with other tools meant to measure anxiety, such as the MASC

(Wood et al., 2002).

Diagnostic Interview Schedule for Children Version 4.0 (DISC 4.0) is a structured

diagnostic interview with a good deal of psychometric evidence (Shaffer, Fisher, Lucas, Dulcan,

& Schwab-Stone, 2000). The DISC 4.0 has evidenced solid test-retest reliability (combined

parent and youth interview κ = .48-.86) when assessing anxiety disorders in youths in

community settings. The DISC 4.0 was used to assess the presence of anxiety diagnoses in the

YAS. The DISC 4.0 is a respondent-based, computer-assisted test that requires respondents to

answer simple yes/no questions. Youth responses were combined with separate parent responses

in order to determine diagnoses. Pairs of condition-blind independent evaluators, constituted of

one research assistant and one clinical psychology graduate student, conducted the DISC 4.0.

Evaluators were trained to adhere exactly to the DISC 4.0 protocol. Interviews were recorded,

randomly sampled, and reviewed by study supervisors throughout the study to assess adherence.

More information regarding the use of the DISC 4.0 in the YAS can be found in Southam-Gerow

et al. (2010), and more specific information regarding the DISC 4.0 can be found in Shaffer et al.

(1996).

Assessment Procedures

Kendall et al. (2008) trial. Participants were recruited through various community

referral sources (e.g., community clinics and practitioners, school systems, and flyers). After

being contacted by a parent, Temple University Child and Adolescent Anxiety Disorders Clinic

staff performed a brief phone screen with the parent to assess a participant's eligibility to be

included in the trial (i.e., youth was age 7 to 14 and the parent was interested in receiving

treatment for their youth). After screening, eligible participants were invited to the clinic, where

independent evaluators conducted separate diagnostic interviews with the parent and the youth. Diagnoses were determined using the ADIS-C/P (Silverman & Albano, 1996). If the parent or youth reported symptoms associated with an anxiety diagnosis, and the symptoms were deemed sufficiently severe by the clinician, the youth was assigned a principal anxiety diagnosis. It was possible to have multiple principal anxiety disorders, but the principal disorder was determined by the severity of the CSR on the integrated youth and parent report. Participants were excluded from the study if the youth had intellectually disability, suffered from a disabling medical condition, was receiving concurrent treatment, was taking antidepressant or antianxiety medication, was experiencing psychotic symptoms, or if neither of the youth's parents spoke English.

If a youth met inclusion criteria for the study, the youth was randomly assigned to condition. Restricted randomization was used to balance participants in each condition. Cases were randomly assigned to therapists. While all therapists were trained to use the three treatment modalities, each therapist delivered only one modality in this trial. The ADIS-C/P, CBCL-A and other outcome instruments were administered before treatment started (Time 1; T1), post-treatment (T2), and 1-year follow up (T3). All procedures were approved by the Temple University Internal Review Board (Kendall et al., 2008).

YAS. Participants in the YAS were clinically-referred youths from six community mental health clinics in the Los Angeles area. Participants and their parents were given the opportunity to participate in this study during routine intakes at six public, urban community mental health clinics. After the intake session, a diagnostic interview was conducted in order to identify principal anxiety diagnoses. In the YAS, diagnoses were determined using the DISC 4.0 (Schaffer et al., 1996). Symptom counts from parent and youth reports were combined. The

project and clinic staff, as well as the parents, met to discuss diagnostic feedback, symptoms and

functioning, referral problem, and whether or not the identified principal anxiety diagnosis was a

priority in treatment. Participants were considered for inclusion in the trial if the youth (a) was

aged 8 to 15, (b) met criteria for an anxiety disorder (GAD, SAD, SOP, or Specific Phobia [SP],

and (c) if anxiety disorders were considered the primary treatment target. Participants were

excluded if they suffered from pervasive developmental disorder or intellectual disability.

If a youth met inclusion criteria, the youth was assigned to either UC or ICBT using

block randomization. Block randomization was used to balance conditions based upon variables

such as sex, the clinic in which the youth was receiving treatment, and language that the youth

speaks. Therapists were also randomly assigned to UC or ICBT using block randomization.

Independent evaluators were blind to condition and performed diagnostic assessments before

treatment began (T1) and at the conclusion of treatment (T2) for both the ICBT and UC

conditions. The assessment procedures were standardized and supervised throughout the

assessment schedule. To assess internalizing symptoms, the CBCL was administered at the same

time as the diagnostic interviews were conducted. All procedures were approved by IRB

(Southam-Gerow et al., 2010).

Observational Procedures

Coders. Two doctoral students used the TPOCS-RS to code recordings from both

studies. Both students were female (one Asian-American and one Caucasian) and were, on

average, 28.60 years of age (*SD* = 1.70). Coders used the TPOCS-RS to characterize the

therapeutic interventions being delivered throughout each session. A three-person team (two

clinical psychology doctoral students and one licensed clinical psychologist) coded the TPOCS-

A to characterize the alliance between therapist and youth in each session. Coders were age 30-

32 years (*M* = 30.67*, SD* = 1.15*).* Coders for both tools were blind to treatment condition and all

coding was randomly assigned.

Coder training. Coders underwent a rigorous training regimen. Training consisted of

reading scoring manuals that described each item (definition of the item, examples of the item,

and what the item is and is not), coding practice sessions, and attending coding meetings that

consisted of discussion of items and coding discrepancies across coders. To be considered

certified (i.e., competent enough to conduct coding for the study), it was necessary for coders to

attain adequate ICCs (>.60; Cicchetti, 1994) over 32 sessions that were master coded by the

study Principal Investigators. Session recordings were randomly assigned to coders. Regular

meetings and ICC assessment were conducted to prevent coder drift (McLeod et al., 2015).

Sampling of treatment sessions. Data from all available and codeable session recordings

were used with the following exceptions. First, the first and last session were excluded because

those sessions likely contained content from intake and termination sessions, which do not

necessarily focus upon the delivery of therapeutic interventions. Second, it was expected that two

sessions of the Coping Cat treatment would include parents, and because this study focused upon

assessing only the effects of therapeutic interventions intended to be delivered to youths, data

from the sessions that did not include the youth were excluded by eliminating those where coders

on the TPOCS-RS indicated the presence of the parent during session. Finally, recordings were

excluded if a recording was (1) missing or damaged, (2) shorter than 15 minutes, (3) contained

less than 15 audible minutes, or (4) less than 75% of the session was in English.

Scoring of treatment sessions. All qualifying sessions were double coded. Item scores

were calculated for each individual session by averaging coders' scores together. After this, the

items in the respective TPOCS-RS subscale were summed and averaged, creating a session-by-

session subscale score for the Coping Cat, Family, and Psychodynamic subscales. Finally, all subscale session scores were averaged across treatment for each individual client. For example, if a client was seen for seven sessions, three total scores represented the extent to which Coping Cat, Family, and Psychodynamic interventions were delivered across the seven sessions.

Data Analytic Strategy

Preliminary analyses

Preliminary analyses consisted of (1) a comparison of current youth and therapist subsamples to those of the parent subsamples, (2) a comparison of the youth and therapist samples (YAS versus Kendall) with regard to demographic and treatment characteristics, (3) assessment and handling of missing data, and (4) dependent variable effect size calculation. All analyses were conducted in SPSS 23 (IBM Corp., 2015).

Subsample to parent sample comparisons. The purpose of these analyses was to identify whether the current subsamples ($n = 51$ for Kendall, $n = 17$ for YAS) were representative of the samples from which they were pulled ($n = 55$ for Kendall, $n = 24$ for YAS). These analyses were calculated to inform later interpretations of results by assessing for sample bias. Only the ICBT condition of each study was included in the analyses. For all comparisons of continuous variables, analyses of variance (i.e., F tests) were used. For all comparisons of categorical variables, chi-square tests were used.

Kendall subsample versus parent sample. Subsample to parent sample analyses were conducted to assess the extent to which the Kendall subsample differed from the parent sample of youths. The following continuous variables were compared: client age, CBCL-A pre- and post-treatment, MASC scores pre- and post-treatment, total number of anxiety diagnoses, total number of externalizing diagnoses, weeks in treatment, and number of sessions held. In addition,

the following categorical variables were compared using chi-square tests: client sex, client ethnicity, client family annual income, therapist sex, and therapist ethnicity.

YAS subsample versus parent sample. These analyses were conducted to assess the extent to which the YAS subsample differed from the parent sample. The following continuous variables were compared: client age, YAS youth and parent factor scores, weeks in treatment, number of sessions attended, and counts of composite pre- and post-treatment diagnoses. The following categorical comparisons were conducted: client sex, client race/ethnicity, and annual family income.

Subsample to subsample comparisons. These analyses consisted of comparing demographic and clinical variables of both client and therapist subsamples for the current study to one another (i.e., YAS subsample versus Kendall subsample). This was done in order to identify if any findings could be related to between-sample demographic differences. The following continuous variables were compared: client age, CBCL-Total (pre- and post-treatment), CBCL-Internalizing (pre- and post-treatment), CBCL-Externalizing (pre- and post-treatment), CBCL-A (pre- and post-treatment), TPOCS-RS subscales (Coping Cat, Psychodynamic, and Family), number of sessions, and weeks in treatment. MASC scores were not compared because the MASC was not collected in YAS. The following categorical variables were compared: client sex, client race/ethnicity, and client family annual income.

Missing data. Missing data for predictor and outcome variables included in the primary analyses were analyzed. Little's Missing Completely at Random (MCAR) test is commonly used to assess the extent to which data are missing systematically or at random (Little, 1988). All quantitative variables to be analyzed were entered into Little's MCAR test. Kendall and YAS analyses were conducted separately because of differences in collected instrumentation and

36

conceptualization of outcomes, namely that only youths in the YAS study did not complete the MASC and youths in the Kendall study did not have an associated parent or youth factor score.

Missing data can be handled in different ways. Some methods (e.g., listwise deletion) remove a participant from the sample if they are missing data from any variable, while other methods (e.g., imputation) attempt to fill in the missing data by producing scores for missing data based upon various methods (Schafer & Graham, 2002). One type of imputation, known as Last Observation Carried Forward (LOCF), is dependent on data being missing completely at random, and replaces missing post-treatment data with the youth's pre-treatment score on the same variable. For both YAS and Kendall, if post-treatment scores were missing, a LOCF approach was used primarily to ensure the maximum sample. The LOCF-included database was used for both preliminary and primary analyses, with the exception of Little's MCAR test.

Effect size calculation. Effect sizes are commonly used as an indicator of clinical significance to describe the magnitude of change between two groups or a pre- and post-treatment outcome (Ferguson, 2009). Cohen's *d* was calculated using Morris and DeShon's (2002) equation, which corrects for the dependence between group means, and thus is ideal for within-group (i.e., repeated measures) effect size calculation. In Kendall, effect sizes for the CBCL-A and MASC were calculated. In YAS, effect sizes for the YAS parent and youth factor scores were calculated. To reduce the amount of statistical analyses performed in YAS, authors created factor scores for parent- and youth-report that combined multiple instruments. The parent factor score consisted of all parent-reported scores related to symptoms of anxiety, including the STAIC-T-Parent (STAIC-T-P), DISC-P anxiety symptom count, and the CBCL-Anxious-Depressed subscale. The YAS youth factor score consisted of STAIC-T scores and DISC-C anxiety symptom counts. If an effect above a small ($d = .20$; Cohen, 2013) was found, outcome

variables were included in subsequent analyses, as this value is commonly used as a cutoff for practical or clinical significance (Ferguson, 2009).

Primary analyses

Due to differences in sample size, outcome measurement, and methodology, data for Kendall and YAS were initially analyzed separately. If the results of these analyses indicated a relation between treatment delivery and client outcomes, then the data from the two trials would be analyzed together to see if the pattern of results was the same across research (i.e., Kendall) and community (i.e., YAS) settings. Despite being analyzed separately, the two sets of hypotheses followed the same procedure for analyses unless explicitly noted otherwise. All primary analyses were conducted in SPSS 23 (IBM Corp., 2015).

Continuous outcomes. Multiple hierarchical regression analyses were used to assess the relation between prescribed and non-prescribed therapeutic interventions (treatment delivery predictor variables) as measured by the TPOCS-RS (Coping Cat, Family, and Psychodynamic subscales, respectively), and client symptom outcomes (dependent variables) as measured by the CBCL-A and MASC in Kendall and the YAS parent and youth factor scores in YAS. Multiple regression is a commonly-used statistical technique for assessing the relation between a single continuous outcome and one or more predictor variables (Jaccard et al., 2006). These analyses used residualized change scores (Manning & DuBois, 1962) for all outcomes to control for pre-treatment severity. Unlike simple change scores (i.e., taking the simple difference between pre- and post-treatment scores), residualized change scores are not correlated with the pre-treatment score, and thus reduce the possibility that any observed relation is a function of a youth's magnitude of change based just upon pre-treatment scores. Residualized change scores have been shown to be more reliable than simple change scores or base-free (Tucker, Damarin, &

38

Messick, 1966) change scores (Williams, Zimmerman, & Mazzagatti, 1987). In practice, residualized change scores account for the outcome score change from pre-treatment to post-treatment by controlling for the pre-treatment outcome score in step 1 of the regression model.

Hypothesis 1. Hypothesis 1a posited that the delivery of prescribed interventions (i.e., TPOCS-RS Coping Cat subscale) would correspond to lower anxiety symptoms (CBCL-A and MASC scores), while hypothesis 1b posited that the delivery of prescribed interventions (i.e., TPOCS-RS Coping Cat subscale) would correspond to better client outcomes over and above the effects of non-prescribed therapeutic interventions (i.e., TPOCS-RS Family and Psychodynamic subscales).

Hypothesis 2. Hypothesis 2a posited that the delivery of prescribed interventions would correspond to lower anxiety symptoms (parent and youth factor scores), while hypothesis 2b posited that the delivery of prescribed interventions (i.e., TPOCS-RS Coping Cat subscale) would correspond to lower anxiety symptoms (parent and youth factors scores) over and above the effects of non-prescribed therapeutic interventions (i.e., TPOCS-RS Family and Psychodynamic subscales).

Assumption checking. Recommendations of Tabachnik and Fidell (2007) were used to guide assumption checking and model building. Hierarchical multiple regression analyses assume a number of characteristics related to the analyzed data. These data were checked to ensure that the data met assumptions of hierarchical multiple regression and to maximize the confidence with which findings were interpreted. Thus, normality, linearity, and other information regarding assumptions (detailed below) were reported, as well as potential problems with those data relative to the assumptions.

First, multiple hierarchical regression assumes that variables are normally distributed.

According to Osbourne and Waters (2002), skewness and kurtosis can affect significance tests and the observed relation between predictor and dependent variables (e.g., increase the probability of type I error) in multiple regressions. Skewness and kurtosis of predictor variables and client symptom outcomes were investigated and reported.

The presence of univariate and multivariate outliers can also affect normality. Thus, univariate outliers were assessed by converting scores into standard (i.e., z) scores. It was assumed that 5% of the sample would have standard values greater/less than $z = \pm 1.96$, 1% would have standard values greater/less than $z = \pm 2.58$, and .001% would have standard values greater/less than $z = \pm 3.29$ After assessment of univariate outliers, multivariate outliers (i.e., cases outside the expected range on more than one variable) were assessed by using the Mahalanobis distance test, which assigns a value to each youth indicating if that youth is a multivariate outlier (i.e., outlier on a combination of variables). Information on univariate and multivariate outliers was reported.

The second major assumption of multiple hierarchical regression is that variables are linearly related to dependent variables. Curvilinear relations between predictor and dependent variables that are analyzed linearly may produce statistics that underestimate the relation between the variables and increase the possibility of Type II error (Osbourne & Waters, 2002). As such, variables were assessed for linearity by creating a scatter plot and ensuring that a curvilinear relation between the predicting variables and the dependent variables did not exist. Variables were also assessed for homoscedasticity (i.e., that variability in scores from one variable will be similar at all values of other variables) by using a residuals scatterplot and assessing whether residuals were equal at all values of the dependent variable.

Finally, hierarchical multiple regression assumes that no two predictor variables are

40

highly correlated (i.e., $r > .70$) with one another, as this can cause analyses to be unstable and because it is difficult to disentangle the effects of predicting variables on dependent variables. Thus, data were checked for multicollinearity (i.e., predicting variables are highly correlated with one another) and singularity (i.e., predicting variables are perfectly correlated with one another) by creating a correlation table and assessing the extent to which predicting variables are correlated with one another.

   Model building. Model building proceeded in the following steps.

   Step 1

   To begin, three models were built for each outcome variable. The purpose of the initial models was to obtain a basic understanding of the nature of the relation between the individual treatment delivery predictors and client symptom outcome. Because later hypotheses (1b and 2b, specifically) were dependent upon first establishing the nature of the relation between the TPOCS-RS Coping Cat subscale and symptom outcomes, this information was to be used to inform later models that included all predictor variables and covariates, if those models were necessary to create. Each model contained one predictor variable (TPOCS-RS Coping Cat, Family, or Psychodynamic subscale) and the residualized change score for one outcome (CBCL-A and MASC in Kendall; YAS parent and youth factor in YAS). Thus, post-treatment outcome score was entered as the dependent variable, the pre-treatment outcome score was entered in step 1 of the regression, and the predictor was entered in step two. The relation between treatment delivery and outcome in these analyses was reported.

   Step 2

   If no relation was observed between the TPOCS-RS Coping Cat subscale and outcome variables, then hypothesis 1a and 2a were disproved. Thus, hypotheses 1b and 2b were not

41

explored further because the underlying premise of hypotheses 1b and 2b relied on a significant relation between the TPOCS-RS Coping Cat subscale and the outcome.

Step 3

If a relation was found specifically between TPOCS-RS Coping Cat and client symptom outcome variables, covariates and predictors were entered in a hierarchical fashion into models. The purpose of this step was to explore, in depth, the relation between the treatment delivery variables and outcome variables. The dependent variable entered into the model was the relevant parent- or youth-symptom outcome variable, as measured by the CBCL-A and MASC in Kendall and the YAS parent and youth factor scores, respectively. Covariates were entered in step 1 to control for their effects on symptomology at post-treatment. In step 1, the respective pre-treatment score was entered to control for pre-treatment symptomology. Second, the alliance was entered, as it has been shown to be related to the treatment delivery and client symptom outcomes (Barber et al., 2006). In step 2, TPOCS-RS Family and Psychodynamic subscales were entered to assess for the effect of non-prescribed therapeutic interventions on symptomology. In step 3, the TPOCS-RS Coping Cat subscale was entered to assess for the effect of prescribed therapeutic interventions on symptomology over and above the effect of non-prescribed interventions.

Categorical outcomes. Hierarchical binomial logistic regression analyses were used. The purpose of these analyses was to assess the relation between prescribed and non-prescribed therapeutic interventions (predictor variables), as measured with the TPOCS-RS (Coping Cat, Family, and Psychodynamic subscales), and remission of primary diagnosis (dependent variables), as measured by the ADIS-C/P in Kendall and the DISC 4.0 in YAS. Hierarchical binomial logistic regression analyses are commonly used in clinical psychology in order to

perform analyses with one or more predicting variables and a categorical outcome (Tabachnik & Fidell, 2007).

Hypothesis 3. Hypothesis 3a posited that the delivery of prescribed interventions (i.e., TPOCS-RS Coping Cat subscale) would correspond to a higher chance of diagnostic remission of ADIS-C/P primary diagnosis in the Kendall trial, while hypothesis 3b posited that the delivery of prescribed interventions would correspond to changes in client outcomes over and above the effects of non-prescribed therapeutic interventions (i.e., TPOCS-RS Family and Psychodynamic subscales).

Hypothesis 4. Hypothesis 4a posited that the delivery of prescribed interventions (i.e., TPOCS-RS Coping Cat subscale) would correspond to a higher chance of diagnostic remission of DISC 4.0 primary diagnosis in the YAS trial, while hypothesis 4b posited that the delivery of prescribed interventions would correspond to changes in client outcomes over and above the effects of non-prescribed therapeutic interventions (i.e., TPOCS-RS Family and Psychodynamic subscales).

Assumption checking. Hierarchical binomial logistic regression assumes four main characteristics of the analyzed data. First, the dependent variable of interest must be a dichotomous variable that is coded such that the value 1 represents the desired outcome. This outcome must contain all categories of the variable. In addition, no youth should be in both categories at one time (i.e., categories are mutually exhaustive). Second, these analyses necessitate one or more continuous or categorical predictor variables. Third, the youths must be independent of one another. Fourth, a linear relation must be observed between the continuous predictor variable(s) and the logit transformation of the dependent variable. This final assumption was checked by using the Box-Tidwell test for linearity. To perform the Box-Tidwell

43

test, the interaction between the predicting variable and the predicting variable's natural log was tested to assess whether the interaction was significantly related to the dependent variable. If the test was non-significant, model building proceeded.

Model building

The steps of model building process were the same as model building for continuous outcomes. The only major difference in the models was that because the outcome variable in logistic regression analyses was categorical and not continuous, no residualized change score was used. Thus, in these models, the predictor variables were entered in step 1 as opposed to step 2.

Chapter 4

Results

Preliminary Analyses

      Subsample to parent sample comparisons. The parent sample and current Kendall

subsample did not differ significantly on any analyzed variables. The current YAS subsample

differed significantly from the parent sample, such that the current subsample was in treatment

for more weeks, attended more sessions, had lower youth-report STAIC-Trait scores pre-

treatment, and had higher counts of composite pre-treatment diagnoses.

      Subsample to subsample comparisons. With regard to continuous variables, significant

differences (as measured by the *F* statistic) were observed for the following variables: TPOCS-

RS Coping Cat subscale, TPOCS-RS Family subscale, TPOCS-RS Psychodynamic subscales,

number of sessions, and weeks in treatment. With regard to categorical variables, significant

differences were observed for client sex, client ethnicity, and annual family income. See Table 2

for results.

      Missing data. With regard to predictor variables (Coping Cat, Psychodynamic, and

Family subscales), a total of 1,098 total sessions were held, and 744 (67.7%) were coded (66.0%

Kendall, $n = 532$; 75.0% YAS, $n = 212$). T-tests showed no significant difference in the overall

percentage of coded sessions $t(66) = 1.85$, $p = .07$ between the Kendall and YAS samples. There

also was no significant difference with regard to sessions not coded as a function of being in the

first (67.7%) or second (67.9%) half of treatment, $t(67) = -.07$, $p = .95$. For pre-treatment

symptom outcome data in Kendall, missing data for TPOCS-RS subscales ranged from 0-2.0%.

For post-treatment symptom outcome data in Kendall, missing data for TPOCS-RS subscales

ranged from 2.0-14.0%. No pre-treatment factor scores (parent or youth) in YAS were missing,

while 29.0% ($n = 5$) post-treatment parent and 29.0% ($n = 5$) post-treatment youth factor scores were missing. Little's MCAR tests indicated that values analyzed in Kendall ($\chi^2 = 41.26$, DF = 52, $p = .86$) and YAS ($\chi^2 = 25.42$, DF = 23, $p = .33$) were missing completely at random. The LOCF approach was implemented.

Effect size calculation. Effect sizes for outcome variables in Kendall and YAS from pre- to post-treatment ranged from small to medium (.45 – .76). This included all dependent variables from pre- to post- treatment in all outcome variables in both samples. See Table 3 for results.

Primary analyses

Continuous outcomes. The following information corresponds to models that assessed the relation between treatment delivery and continuous symptom outcomes in Kendall and YAS.

Treatment delivery and parent-reported symptomology in Kendall.

Assumption checking. The distributions of the CBCL-A and two predictor variables (TPOCS-RS Coping Cat and Family subscales) were normal, while the TPOCS-RS Psychodynamic subscale was not, with skewness of 3.06 (*SE* = .33) and kurtosis of 10.78 (*SE* = .66). No univariate or multivariate outliers were detected. In accordance with step 1 of the model building plan, no transformations or alterations to the data were made for initial models. All other assumptions of multiple hierarchical regression were met. First, the assumption of linearity was met, as no curvilinear relations were found after scatterplots were created. Second, the assumption of homoscedasticity was met, as no heteroscedasticity was found after residual scatterplots were made. Third, the assumption of multicollinearity was met, as no correlations between predictor variables above .70 were observed.

Table 2

Youth Descriptive Data and Group Comparisons for Kendall and YAS subsamples

Variable	M *(SD)* or %		*F* or Chi Square
	Kendall (N = 51)	YAS (N = 17)	
Age	10.36 *(1.90)*	11.32 *(2.32)*	1.17
Sex			
Male	60.80	29.40	5.04*
Race/Ethnicity			18.97**
Caucasian	86.30[a]	41.20	
African-American	9.80	-	
Latino	2.00	17.60[b]	
Mixed/Other	2.00	5.90	
Not Reported	-	35.30[b]	
CBCL			
Total (pre)	63.18 *(8.44)*	64.19 *(7.34)*	.47
Total (post)	58.08 *(9.20)*	59.94 *(7.45)*	.74
Internalizing (pre)	67.40 *(8.37)*	66.38 *(8.33)*	.04
Internalizing (post)	62.22 *(7.96)*	62.18 *(8.20)*	.01
Anxiety (pre)	14.30 *(5.94)*	9.75 *(4.95)*	.89
Anxiety (post)	10.78 *(5.36)*	6.47 *(4.20)*	.54
TPOCS-RS subscales			
Coping Cat	3.42 *(.50)*[a]	2.36 *(.70)*	6.13*
Psychodynamic	1.07 *(.11)*	1.26 *(.27)*[b]	40.12***
Family	1.14 *(.13)*	1.43 *(.30)*[b]	20.32***
Primary Diagnoses			22.81**
GAD	37.30[a]	5.90	
SAD	29.40	35.30	
SOP	33.30	23.50	
SP	-	35.30[b]	
Family Income			7.92**
Up to 60k per year	35.30	70.60[b]	
Number of Sessions	15.92 *(1.43)*	16.82 *(5.02)*[b]	19.78***
Weeks in Treatment	19.52 *(3.97)*	26.38 *(10.41)*[b]	23.35***

Note. Kendall = Individual Cognitive Behavioral Therapy in Kendall et al. (2008) study, *YAS* = Individual Cognitive Behavioral Therapy in Southam-Gerow et al. (2010). CBCL = Child Behavior Checklist, GAD = generalized anxiety disorder, SAD = separation anxiety disorder, SOP = social phobia, SP = specific phobia, TPOCS-RS = Therapy Process Observational Coding System – Revised Strategies Scale.
[a] =Kendall > YAS
[b] = YAS > Kendall
*p < .05, **p < .01, ***p < .001

Model building. A one unit increase in TPOCS-RS Coping Cat scores was associated

with a 1.14 unit decrease in parent-reported symptomology, but this relation was not significant,

β = -.11, $t(47)$ = -1.00, p = .32. A one unit increase in TPOCS-RS Family subscale scores

yielded a 1.40 unit decrease in parent-reported symptomology, but this association was not

significant, β = -.03, $t(47)$ = -.30, p = .77. A one unit increase in TPOCS-RS Psychodynamic

subscale was associated with an a 6.64 unit increase in CBCL-A scores, but this relation was not

significant, β = .13, $t(47)$ = 1.20, p = .24. In addition, models that contained TPOCS-RS

predictors did not explain variance in CBCL-A scores significantly better than models with only

the baseline CBCL-A measurement. Because no predictors were related to outcomes, model

building did not proceed. See Table 4 for more information.

Table 3

Pre- to post-treatment outcome score effect sizes

Instrument	Pre-treatment M*(SD)*	Post-treatment M*(SD)*	*r*	Cohen's *d*
MASC	49.65 *(21.72)*	40.67 *(20.55)*	.56	.45*
CBCL-A	14.30 *(5.94)*	10.78 *(5.36)*	.66	.76**
YAS parent factor	0.02 *(.90)*	-0.64 *(.82)*	.52	.78**
YAS youth factor	0.10 *(.76)*	-0.24 *(1.05)*	.30	.32*

Note. Analyses of MASC and CBCL-A included youths in Kendall sample only. Analyses of
YAS parent and youth factor scores included youths in YAS sample only.
Based upon Cohen (2013), *small effect size (.20), **medium effect size (.50), ***large effect
size (.80)

Treatment delivery and youth-reported symptomology in Kendall.

Assumption checking. The MASC was normally distributed. Thus, no alterations or

transformations were required for this variable. One multivariate outlier was identified in the

TPOCS-RS Psychodynamic model with the MASC outcome (χ^2 = 23.23), but no changes were

made to the data for initial models. No other univariate or multivariate outliers were detected. All

other assumptions of multiple hierarchical regression were met.

Model building.

A one unit increase in TPOCS-RS Coping Cat scores was associated with a 1.40 unit

increase in MASC scores, but this relation was not significant, $\beta = .03$, $t(48) = 2.80$, $p = .78$. A

one unit increase in TPOCS-RS Family subscale was associated with a decrease of 13.55 units in

MASC scores, though this relation was not significant, $\beta = -.08$, $t(48) = -.69$, $p = .49$. The

models containing the TPOCS-RS Coping Cat and Family subscales did not explain more

variation than models that contained only the baseline measurement. The initial model for

TPOCS-RS Psychodynamic suggested that the subscale was significantly related to an increase

in MASC scores, $\beta = .31$, $t(48) = 2.80$, $p = .007$, such that a one unit increase in this subscale

was related to an increase of 60.82 units in MASC scores. See Table 4 for information on initial

models.

As noted above, outliers and variables that do not conform to a normal distribution may

affect regression analyses. As such, efforts were made to normalize the distribution of non-

normal variables. The purpose of this was to ensure that the relation between TPOCS-RS

Psychodynamic scale and MASC scores was not a function of problems with the normality of

variables in the analyses. This was first done by addressing univariate and multivariate outliers.

If it was determined that univariate outliers were the cause of non-normality, values outside of 3

standard deviations were winsorized (i.e., pulled closer to the mean), allowing for the analyses to

respect the score's considerable difference from the mean while maintaining the integrity of

analyses. This was done by converting the score into the equivalent of a standard score of 3.29 or

-3.29, or 3 standard scores away from the mean. Critical chi-square values were used to identify

the critical value for a multivariate outlier. If a case (i.e., youth) obtained a score greater than 16.27 on the Mahalanobis distance test, the youth was eliminated from the analysis.

Data transformations (e.g., natural log, square root) are often used in psychosocial research to make data more closely resemble a normal, bell-shaped distribution (Feng et al., 2014). However, transformations of data can significantly alter the interpretability of data, making hypothesis testing and interpretation of statistical analyses more difficult (Feng et al., 2014; Osborne, 2010). To maximize interpretability of results, data transformations were not used as part of the data transformation process.

No univariate outliers were identified in the sample. One multivariate outlier was identified ($\chi^2 = 23.23$). After this case was removed, the analyses were subsequently re-conducted. After re-analysis, the TPOCS-RS Psychodynamic subscale was not significantly related to MASC scores, $\beta = .22$, $t(47) = 1.78$, $p = .08$. In addition, the model that included the MASC did not explain more variance in post-treatment MASC scores than the model which contained only the baseline measurement, $\Delta R^2 = .05$, $p = .08$.

Treatment delivery and parent-reported symptomology in YAS

Assumption checking. The distributions of the YAS parent factor and all predicting variables (TPOCS-RS Coping Cat, Family, and Psychodynamic subscales) were normally distributed. In addition, no univariate or multivariate outliers existed. All other assumptions of multiple hierarchical regression were met.

Model building. No TPOCS-RS subscale was significantly related to YAS parent factor scores. A one unit increase in TPOCS-RS Coping Cat scores was associated with a .13 increase in parent-reported symptomology, but this relation was not significant, $\beta = .11$, $t(16) = .49$, $p = .63$. A one unit increase in TPOCS-RS Family subscale was associated with .48 unit decrease

50

in parent-reported symptomology, but this relation was not significant, $\beta = -.18$, $t(16) = -.77$, $p = .45$. The TPOCS-RS Psychodynamic subscale yielded a .16 unit decrease in parent-reported symptomology, but this association was not significant, $\beta = -.05$, $t(16) = -.22$, $p = .83$. Models that contained TPOCS-RS predictors did not explain variance in YAS parent factor scores significantly better than models with only the baseline YAS parent factor score measurement. Because no predictors were related to outcomes, model building did not proceed. See Table 5 for more information.

Table 4

Relation between treatment delivery and symptom outcome in Kendall

Predictors	CBCL-A					MASC				
	β	$p <$	R^2	ΔR^2	$p <$	β	$p <$	R^2	ΔR^2	$p <$
Coping Cat subscale										
Step 1										
Baseline measurement	.60	.00***	.44			.53	.00***	.31		
Step 2										
Coping Cat	-.11	*ns*	.45	.01	*ns*	.03	*ns*	.31	.00	*ns*
Family subscale										
Step 1										
Baseline measurement	.60	.00***	.44			.53	.00***	.31		
Step 2										
Family	-.03	*ns*	.44	.00	*ns*	-.08	*ns*	.32	.01	*ns*
Psychodynamic subscale										
Step 1										
Baseline measurement	.60	.00***	.44			.53	.00***	.31		
Step 2										
Psychodynamic	.13	*ns*	.46	.02	*ns*	.31	.00**	.41	1.0	.00**

Note. ns = not significant. CBCL-A = Child Behavior Checklist – Anxiety subscale. MASC = Multidimensional Anxiety Scale for Children. Coping Cat Subscale = TPOCS-RS Coping Cat Subscale. Family subscale = TPOCS-RS Family subscale. Psychodynamic subscale = TPOCS-RS Family subscale.
*p < .05, **p < .01, ***p < .001

Treatment delivery and youth-reported symptomology in YAS.

Assumption checking. The YAS youth factor and all predicting variables (TPOCS-RS Coping Cat, Family, and Psychodynamic subscales) were normally distributed. In addition, no univariate or multivariate outliers existed. All other assumptions of multiple hierarchical regression were met.

Model building. The TPOCS-RS Coping Cat, Family, and Psychodynamic subscales were not significantly related to YAS youth factor scores. A one unit change TPOCS-RS Coping Cat subscale scores were associated with a decrease of .46 units in YAS youth factor scores, but this relation was not significant, $\beta = -.31$, $t(16) = -1.26$, $p = .23$. A one unit change in TPOCS-RS Family scores was associated with an increase of 1.52 units in YAS youth factor scores, but this relation was not significant, $\beta = .43$, $t(16) = 1.85$, $p = .09$. Finally, a one unit change in TPOCS-RS Psychodynamic scores was associated with a 1.15 unit increase in post-treatment YAS youth factor scores, but the relation was not significant, $\beta = .30$, $t(16) = 1.00$, $p = .33$. Models containing TPOCS-RS subscales did not explain variance in the dependent variable significantly better than models that only contained baseline measurement. See Table 5 for more information.

Categorical outcomes. The following information corresponds to models that assessed the relation between treatment delivery and categorical diagnostic remission outcomes in the Kendall and YAS subsamples.

Treatment delivery and diagnostic status at post-treatment in Kendall.

Assumption checking. The data met all assumptions of hierarchical binomial logistic regression. The Box-Tidwell tests for the TPOCS-RS Coping Cat subscale ($p = .82$), TPOCS-RS

Family subscale ($p = .93$), and TPOCS-RS Psychodynamic subscale ($p = .08$) were not significant, indicating a linear relation between the continuous predictor variable(s) and the logit transformation of the dependent variable.

Model building. No predictor variables were significantly related to post-treatment diagnostic status in Kendall. The model that included the TPOCS-RS Coping Cat subscale did not explain variation in diagnostic remission significantly better than comparison (i.e., constant-only) model, $\chi^2 = .08$, $p = .78$, $df = 1$. The TPOCS-RS Coping Cat subscale was related to increases in the probability of diagnostic remission, but this relation was not significant (Wald $= .08$, $p = .78$). The model that included the TPOCS-RS Family subscale also did not explain variation in diagnostic remission significantly better than the comparison model $\chi^2 = 3.22$, $p = .07$, $df = 1$. The TPOCS-RS Family subscale was related to decreases in the probability of diagnostic remission, but this relation was not significant (Wald $= 2.94$, $p = .09$). Finally, the model that included the TPOCS-RS Psychodynamic subscale did not explain variation in diagnostic remission significantly better than the comparison model $\chi^2 = .02$, $p = .90$, $df = 1$. The TPOCS-RS Psychodynamic subscale was related to decreases in the probability of diagnostic remission, but this relation was not significant (Wald $= .02$, $p = .90$). See Table 6 for more information.

Treatment delivery and diagnostic status at post-treatment in YAS

Assumption checking. The data met all assumptions of hierarchical binomial logistic regression. The Box-Tidwell tests for the TPOCS-RS Coping Cat subscale ($p = .91$), TPOCS-RS Family subscale ($p = .99$), and TPOCS-RS Psychodynamic subscale ($p = .84$) were not significant, indicating a linear relation between the continuous predictor variable(s) and the logit transformation of the dependent variable.

Table 5

Relation between treatment delivery and symptom outcome in YAS

Predictors	YAS parent factor					YAS youth factor				
	β	$p <$	R^2	ΔR^2	$p <$	β	$p <$	R^2	ΔR^2	$p <$
Coping Cat subscale										
Step 1										
Baseline measurement	.52	.03*	.27			.30	ns	.09		
Step 2										
Coping Cat	.11	ns	.29	.01	ns	-.31	ns	.18	.09	ns
Family subscale										
Step 1										
Baseline measurement	.52	.03*	.27			.30	ns	.09		
Step 2										
Family	-.18	ns	.30	.03	ns	.43	ns	.27	.18	ns
Psychodynamic subscale										
Step 1										
Baseline measurement	.52	.03*	.27			.30	ns	.09		
Step 2										
Psychodynamic	-.05	ns	.28	.00	ns	.30	ns	.15	.06	ns

Note. ns = not significant. Coping Cat Subscale = TPOCS-RS Coping Cat Subscale. Family subscale = TPOCS-RS Family subscale. Psychodynamic subscale = TPOCS-RS Family subscale.
*p < .05, **p < .01, ***p < .001

Model building. No predictor variables were significantly related to post-treatment diagnostic status in YAS. The model that included the TPOCS-RS Coping Cat subscale did not explain variation in diagnostic remission significantly better than comparison (i.e., constant-only model), $\chi^2 = .01$, $p = .92$, DF = 1. The TPOCS-RS Coping Cat subscale was related to increases in the probability of diagnostic remission, but this relation was not significant (Wald = .01, $p = .92$). The model that included the TPOCS-RS Family subscale also did not explain variation in diagnostic remission significantly better than the comparison model $\chi^2 = .01$, $p = .92$, DF = 1.

The TPOCS-RS Family subscale was related to decreases in the probability of diagnostic

remission, but this relation was not significant (Wald = .01, p = .92). Finally, the model that

included the TPOCS-RS Psychodynamic subscale did not explain variation in diagnostic

remission significantly better than the comparison model χ^2 = .02, p = .89, DF = 1. The TPOCS-

RS Psychodynamic subscale was related to decreases in the probability of diagnostic remission,

but this relation was not significant (Wald = .02, p = .89). See Table 6 for more information.

Table 6

The relation between treatment delivery and diagnostic outcome

Predictors	Kendall					YAS				
	B	Odds Ratio	95% CI Lower	Upper	$p <$	B	Odds Ratio	95% CI Lower	Upper	$p <$
Coping Cat subscale	.16	1.17	.39	3.54	*ns*	.07	1.07	.26	4.42	*ns*
Family subscale	-4.50	.01	.00	1.91	*ns*	-.18	.84	.03	23.26	*ns*
Psychodynamic subscale	-.33	.72	.00	136.82	*ns*	-.27	.77	.02	29.98	*ns*

Note. CI = Confidence Interval. *ns* = not significant. Coping Cat Subscale = TPOCS-RS Coping Cat Subscale. Family subscale = TPOCS-RS Family subscale. Psychodynamic subscale = TPOCS-RS Family subscale.
*p < .05, **p < .01, ***p < .001

Chapter 5

Discussion

Understanding the relation between therapeutic interventions and client post-treatment clinical outcomes can help researchers understand how psychosocial treatments work. The current study aimed to add to this literature by assessing the relation between the delivery of an EBT in research and practice settings and client symptom and diagnostic outcomes. This discussion will present interpretations of the study findings and how the findings relate to the literature at large and the conceptual model used to form hypotheses.

What was the nature of the relation between treatment delivery and client outcomes?

The current study found no significant relation between treatment delivery and symptom or diagnostic outcomes in research and practice settings. While youths in both trials showed improvements in outcomes from pre- to post-treatment, very little of this change was explained by what therapeutic interventions were delivered, whether those interventions were prescribed or non-prescribed. The findings related to the delivery of prescribed therapeutic interventions were, for the most part, consistent with past findings (e.g., Webb et al., 2010). Conversely, findings related to the delivery non-prescribed therapeutic interventions are largely new.

The results of the current study present both expected and unexpected findings. With regard to expected findings, the relation between the delivery of more prescribed therapeutic interventions and better outcomes (lower symptom or higher chance of diagnostic remission) trended in the expected direction for four of the six models. In addition, non-prescribed therapeutic interventions were not related to better outcomes in 8 of the 12 models. With regard to unexpected findings, a higher dose of prescribed therapeutic interventions (i.e., higher adherence) was not always related to decreases in symptomology (e.g., CBCL-A models in Kendall sample) and remission of diagnosis. Second, despite the hypothesis that higher doses of

56

prescribed interventions would lead to better outcomes in both settings, prescribed interventions were inconsistently related to symptoms across settings. Overall, the findings were inconsistent within independent variables (e.g., higher doses of prescribed interventions did not always lead to better outcomes), across and within reporter (i.e., youth- and parent- report), across and within outcome type (i.e., diagnostic and symptom), and across setting (i.e., YAS and Kendall).

Two main interpretations can be posited based upon the findings, the study methodology, and analytic plan. First, and most simply, it is possible that delivering prescribed interventions does not affect outcomes. In other words, the therapeutic interventions that are delivered may not be the ingredients of therapy that cause changes in symptoms or diagnostic status. While few, if any studies speak to the relation between non-prescribed interventions and outcomes, a number of studies have shown no relation between prescribed interventions and outcomes (e.g., Becker et al., 2012; Loeb et al., 2005). Webb et al. tackled this possibility in their review, indicating that (1) it is possible that the delivered interventions may not be the most important part of treatment, and thus do not account for significant variance in outcome change from pre- to post-treatment, and (2) the studies that found a significant relation were a product of chance, as evidenced by the fact that the majority of the variance in outcome was due to study differences as opposed to sampling error.

The second interpretation of the results is that the delivery of therapeutic interventions is important in client outcomes, but that the obtained results were a product of weaknesses in the study. Webb et al. (2010) outline a gold standard of methodology and analyses for studying the relation between treatment delivery and outcomes. This study strove to meet those standards but fell short in several ways (e.g., incorporating nesting in statistical analyses). Thus, it is possible that the methodology used to conceptualize and characterize treatment integrity or the analytic

plan used to study the relation between treatment delivery and outcome influenced the extent to which finding a relation between treatment delivery and outcome was possible.

At this time, both interpretations are plausible. Indeed, even if analytic and methodological considerations played a role in the ability of this study to find a relation that does not exclude the possibility that there is no relation. Considering how these findings relate to past evidence and the model used to form hypotheses may be important in understanding how these results came about and also may help researchers conduct studies that paint a more complete picture of the relation between treatment delivery and outcomes.

How does the current study relate to the literature base?

The Webb et al. (2010) meta-analysis and the literature reviewed for the current study have demonstrated a positive relation, negative relation, and no relation between treatment delivery and outcome. One addition to the literature base was the incorporation of non-prescribed therapeutic interventions. Non-prescribed therapeutic interventions have rarely been studied, leaving a gap in the understanding of the way treatment delivery affects client outcomes. In the current study, non-prescribed therapeutic interventions were at times as strongly or more strongly related to outcomes than prescribed therapeutic interventions, indicating that some relation may exist or that these interventions may affect outcomes in some way.

The findings of this study with regard to prescribed interventions echo the discrepant findings from the literature, in that the pattern of results was inconsistent. When reviewing past literature, it appears that the analytic plan in this study may have been oversimplified in two key ways that support the interpretation that analytic flaws contributed to the difficulty in identifying a relation between treatment delivery and outcome. First, the analytic plan in the current study did not use multilevel modeling, which would help account for nesting of sessions within clients

and clients within therapist. Past literature has suggested that variance exists in multiple levels of

therapy, such that delivery of therapeutic interventions varies both within and between therapist

caseloads, meaning that the therapeutic interventions delivered across clients is different with

regard to the therapist (e.g., Imel, Baer, Martino, Ball, & Carroll, 2011). Analyses used in this

study did not take this into account, as linear regression compresses the variance between and

within levels, and thus tends to be less sensitive to the effects of that variance.

Second, the current analytic plan did not take into account the extent to which treatment

delivery may differ over time, and what effects that variation may have on outcomes. Past

research suggests that the extent to which therapeutic interventions are delivered changes over

time (e.g., Owen & Hilsenroth, 2014). However, the use of linear regression does not definitively

exclude the possibility of finding a relation. For instance, some studies that have used linear

regression and correlation (e.g., Goldman & Gregory, 2009; Webb et al., 2012) have found a

significant relation between treatment delivery and outcome. Thus, the limitation of using these

analyses is that they are not as suitable for explaining the data, and thus (1) may decrease the

possibility of finding a relation because analyses are not able to account for variation that may be

responsible for change in the outcome and (2) may not provide a complete picture of the relation

between treatment delivery and outcomes.

How do the findings relate to the conceptual model?

The Doss model indicates that prescribed therapeutic interventions target change

mechanisms that eventually lead to changes in client outcomes. The lack of significant findings

in non-prescribed therapeutic interventions bolsters the model. However, the findings from

prescribed interventions appear contrary to this idea, implying that there is no statistical relation

between the therapeutic interventions meant to target change mechanisms and subsequent change

in client symptom or diagnostic outcomes. Despite finding a trend that appears to be somewhat inconsistent with the model (e.g., increased dose of prescribed interventions led to better and worse outcomes within Kendall), these results do not necessarily mean that the model is invalid or incorrect. Indeed, a number of studies have demonstrated the validity of this model by assessing the relation between specific interventions, change mechanisms, and subsequent changes in outcome, though these studies have not yet connected specific aspects of treatment delivery (i.e., prescribed or non-prescribed therapeutic interventions) in analyses (e.g., Nakamura, Pestle, & Chorpita, 2009; White, Kendrick, & Yardley, 2009). Two potential hypotheses exist: (1) that the interventions measured by the TPOCS-RS and used to characterize prescribed interventions in the present study were not those interventions that led to the change observed in symptom and diagnostic outcomes, or (2) that the use of the TPOCS-RS may have limited the ability to measure the exact therapeutic interventions of Coping Cat due to the fact that the TPOCS-RS items were created to be more general in nature.

Another potential hypothesis, presented by the Doss (2004) model and other researchers (Stiles & Shapiro, 1994) known as the responsiveness critique, attempts to explain the lack of observed findings in some studies that assess the relation between treatment delivery and outcome. The responsiveness critique states that if therapists deliver therapeutic interventions relative to the symptomology and functioning of the client, there should be no linear relation between what interventions were delivered and client outcomes. The responsiveness critique has primarily been discussed in the context of analytic methods such as linear regression and correlation, as these methods are based upon the idea of linear relation between predictor X and outcome Y. The observed pattern of findings, as well as the analytic methods used to study the relation bolster the possibility that the responsiveness critique may have played a role in the

60

study findings. While the responsiveness critique may play a role in the observed findings, it is also possible methodological and analytic study weaknesses were partially responsible.

What were the weaknesses of the current study?

Sample size. The sample size in YAS was limited to 17 youths. Small sample sizes can lead to a lack of power in statistical analyses (Cohen, 1992). However, if the primary limitation of this study was low power, it would be expected that a consistent pattern would emerge from the data in the form of small or medium, statistically insignificant effects. No clear trend emerged. The lack of any clear trend provides some support for the interpretation that the delivery of prescribed interventions is not the most influential predictor of symptom or diagnostic change, though it is still possible that sample size and power played some role.

Multi-level modeling. This study did not use methods that (1) allowed for the consideration of time and (2) accounted for nesting of sessions within youths within therapists, limiting the interpretability of the results. Recent studies have shown that treatment delivery systematically changes over time (e.g., Boswell et al., 2013; Owen & Hilsenroth, 2014) and that measuring and analyzing treatment delivery on a session-by-session basis is helpful in capturing the complexity of those changes (e.g., Webb et al., 2010). While variation in treatment delivery can sometimes be important in finding a relation (e.g., within-case variability in treatment delivery affecting outcome in Owen & Hilsenroth, 2010), that change over time does not necessarily mean that that variation predicts outcomes. For instance, Becker et al. (2012) and Boswell et al. (2013) found significant variation in treatment delivery of CBT, but that treatment delivery was not significantly related to client symptom outcomes. Thus, while this study did not use the analytic method most sensitive to treatment delivery change over time, which could have

decreased the possibility of finding a relation, the study weaknesses do not necessarily undercut the interpretation that there is no relation between treatment delivery and outcome.

What were the strengths of the current study?

The current study had conceptual and methodological strengths that bolstered internal and external validity. First, the TPOCS-RS allows for the measurement of both prescribed and non-prescribed therapeutic interventions. This is a rarely taken, but important step in treatment delivery conceptualization and measurement, as it has implications for understanding the relation between what is delivered and subsequent changes in symptom or diagnostic outcome. The use of an observational tool such as the TPOCS-RS also minimizes the bias presented by other types of treatment integrity measurement strategies (e.g., therapist self-report; DiMatteo, 2004; Hill, 1991; Mowbray et al., 2003). In addition, the TPOCS-RS uses an extensiveness rating scale, which is considered by some to be well-suited to investigating process-outcome relations (Hogue et al., 1996). These methodological strengths (i.e., minimizing bias and the use of extensiveness scores) only indicate that treatment delivery was measured in an optimal way, but does little to bolster the strength of the findings themselves, as the subsequent handling (i.e., making one average session TPOCS-RS score) of treatment delivery data may not have been optimal.

The other primary strength of this study was the use of data from trials in research and practice settings. No reviewed studies, including those in Webb et al., (2010) included data from trials in both settings. This increases the external validity of the study, making the results more generalizable outside the context of this study. However, in much the same way as the TPOCS-RS strengthened methodology but not necessarily the conclusions of the study, the inclusion of both study contexts bolsters the generalizability of the study conclusions insofar as the study conclusions are bolstered by other study factors. In other words, because there is reason to

believe that the internal validity of the study was compromised by an oversimplified analytic plan, the generalizability of the study conclusions suffers as a result.

Future directions

Further investigation into the relation between treatment delivery and client outcomes is warranted. To do this, statistical models should use analyses that allow for the inclusion of time and repeated-measures predictor variables. Multilevel modeling or hierarchical linear modeling are ideal for these types of analyses (Tabachnik & Fidell, 2007). Utilization of these statistical techniques will allow for investigators to identify to what extent treatment delivery changes over time, and how that change affects client outcomes (Kahn, 2011).

Second, few studies measure the delivery of non-prescribed interventions (i.e., treatment differentiation) and incorporate into analyses as a predicting variable. As noted above, including this in measurement efforts and analyses will likely be important when conducting effectiveness trials in community settings, as the delivery of non-prescribed therapeutic interventions is not uncommon (Weisz et al., 2009). Though no statistical relation between treatment differentiation and outcome was found in this study, non-prescribed interventions trended toward significance or affected outcome more strongly than prescribed interventions in some circumstances.

Conclusion

The current study demonstrated no relation between treatment delivery and youth symptom and diagnostic outcomes. These findings, when taken with conceptual and empirical considerations as well as methodological and analytical flaws, should be interpreted and generalized with caution. Implementing solutions to the methodological and analytical considerations presented above may provide researchers with more confidence when studying the relation between treatment delivery and outcome. Despite the array of work that has been

done across EBTs, more work is needed that addresses presented weaknesses and includes

treatment differentiation, thereby forming a more comprehensive idea of how treatment delivery

affects outcomes.

References

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist 4–18 and 1991 Profile.*

Burlington, VT: Department of Psychiatry, University of Vermont.

American Psychological Association Task Force on Evidence-Based Practice for Children and

Adolescents. (2008). *Disseminating evidence-based practice for children & adolescents: A systems approach to enhancing care: APA task force on evidence-based practice with children and adolescent.* Washington, D.C.: American Psychological Association.

Barber, J. P., Gallop, R., Crits-Cristoph, P., Frank, A., Thase, M. E., Weiss, R. D., & Gibbons,

M. B. (2006). The role of therapist adherence, therapist competence, and the alliance in predicting outcome of individual drug counseling: Results from the NIDA Collaborative Cocaine Treatment Study. *Psychotherapy Research*, *16*(2), 229-240. doi:10.1080/10503300500288951

Becker, E. M., Becker, K. D., & Ginsburg, G. S. (2012). Modular cognitive behavioral therapy

for youth with anxiety disorders: A closer look at the use of specific modules and their relation to treatment process and response. *School Mental Health, 4*(4), 243-253. doi:10.1007/s12310-012-9080-2

Bickman, L., & Hamilton, J. D. (2008). A measurement feedback system (MFS) is necessary to

improve mental health outcomes. *Journal of the American Academy of Child & Adolescent Psychiatry, 47*(10), 1114-1119. doi:10.1097/CHI.0b013e3181825af8

Boswell, J. F., Gallagher, M. W., Sauer-Zavala, S. E., Bullis, J., Gorman, J. M., Shear, M. K., . . .

Barlow, D. H. (2013). Patient characteristics and variability in adherence and competence in

cognitive-behavioral therapy for panic disorder. *Journal of Consulting and Clinical Psychology,81*(3), 443-454. doi:10.1037/a0031437

Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology, 52*, 685–716. doi: 10.1146/annurev.psych.52.1.685

Chiu, A. W., McLeod, B. D., Har, K., & Wood, J. J. (2009). Child–therapist alliance and clinical outcomes in cognitive behavioral therapy for child anxiety disorders. *Journal of Child Psychology and Psychiatry, 50*(6), 751-758. doi:10.1111/j.1469-7610.2008.01996.x

Chorpita, B. F., Daleiden, E. L., Ebesutani, C., Young, J., Becker, K. D., Nakamura, B. J., . . . Trent, L. (2011). Evidence-based treatments for children and adolescents: An updated review of indicators of efficacy and effectiveness. *Clinical Psychology: Science and Practice, 18*(2), 154-172. doi:10.1111/j.1468-2850.2011.01247.x

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4)*, 284−290. doi:10.1037/1040-3590.6.4.284

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences.* Academic press.

DiMatteo, M. R. (2004). Variations in patients' adherence to medical recommendations: A quantitative review of 50 years of research. *Medical Care, 42*(3), 200–209. doi:10.1097/01.mlr.0000114908.90348.f9

Doss, B. D. (2004). Changing the way we study change in psychotherapy. *Clinical Psychology: Science and Practice, 11*(4), 368-386. doi:10.1093/clipsy.bph094

Eyberg, S. M., Nelson, M. M., & Boggs, S. R. (2008). Evidence-based psychosocial treatments for children and adolescents with disruptive behavior. *Journal of Clinical Child & Adolescent Psychology, 37*(1), 215-237. doi:10.1080/15374410701820117

Feeley, M., DeRubeis, R. J., & Gelfand, L. A. (1999). The temporal relation of adherence and alliance to symptom change in cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 67*(4), 578-582. doi:10.1037/0022-006X.67.4.578

Feng, C., Wang, H., Lu, N., Chen, T., He, H., Lu, Y., & Tu, X. M. (2014). Log-transformation and its implications for data analysis. *Shanghai Archives of Psychiatry*, *26*(2). doi:10.3969/j.issn.1002-0829.2014.02.009

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice, 40*(5). doi:10.1037/a0015808

Fjermestad, K. W., McLeod, B. D., Heiervang, E. R., Havik, O. E., Öst, L., & Haugland, B. S. (2012). Factor structure and validity of the therapy process observational coding system for child psychotherapy–alliance scale. *Journal of Clinical Child & Adolescent Psychology, 41*(2), 246-254. doi:10.1080/15374416.2012.651999

Goldman, G. A. & Gregory, R. J. (2009). Preliminary relationships between adherence and outcome in dynamic deconstructive psychotherapy. *Psychotherapy Theory, Research, Practice, Training, 46*(4), 480-485. doi:10.1037/a0017947

Gravetter, F., & Wallnau, L. (2014). *Essentials of statistics for the behavioral sciences (8th ed.).* Belmont, CA: Wadsworth.

Hill, C. E., O'Grady, K. E., & Elkin, I. (1992). Applying the Collaborative Study Psychotherapy Rating Scale to rate therapist adherence in cognitive-behavior therapy, interpersonal therapy, and clinical management. *Journal of Consulting and Clinical Psychology, 60*(1), 73-79. doi:0022-006X/92/J3.00

Hoagwood, K., Burns, B. J., Kiser, L., Ringeisen, H., & Schoenwald, S. K. (2014). Evidence-based practice in child and adolescent mental health services. *Psychiatric Services,52*(9), 1179-1189. doi:10.1176/appi.ps.52.9.1179

Hoagwood, K., Jensen, P. S., Petti, T., & Burns, B. J. (1996). Outcomes of mental health care for children and adolescents: I. A comprehensive conceptual model. *Journal of the American Academy of Child & Adolescent Psychiatry, 35*(8), 1055-1063. doi:10.1097/00004583-199608000-00017

Hogue, A., Liddle, H. A., & Rowe, C. (1996). Treatment adherence process research in family therapy: A rationale and some practical guidelines. *Psychotherapy: Theory, Research, Practice, Training, 33*(2), 332-345. doi:10.1037/0033-3204.33.2.332

IBM Corp. Released 2015. IBM SPSS Statistics for Windows, Version 23.0 [Computer software]. Armonk, NY: IBM Corp.

Imel, Z. E. Baer, J. S., Martino, S., Ball, S. A., Carroll, K. M. (2011). Mutual influence in therapist competence and adherence to motivational enhancement therapy. *Drug and Alcohol Dependence, 115*(3), 229-236. doi:10.1016/j.drugalcdep.2010.11.010

Institute of Medicine. (2015). *Improving the quality of health care for mental and substance-use conditions: Quality chasm series.* Washington, DC: National Academies Press.

Iversen, G. (1991). *Contextual Analysis.* Newbury Park, CA: Sage.

Jaccard, J., Guilamo-Ramos, V., Johansson, M., & Bouris, A. (2006). Multiple regression analyses in clinical child and adolescent psychology. *Journal of Clinical Child and Adolescent Psychology, 35*(3), 456-479. doi:10.1207/s15374424jccp3503_11

Judd, C. M., & Kenny, D. A. (1981). Process analysis estimating mediation in treatment evaluations. *Evaluation Review, 5*(5), 602-619. doi:10.1177/0193841X8100500502

Kahn, J. K. (2011). Multilevel modeling: Overview and applications to research in counseling psychology. *Journal of Counseling Psychology, 58*(2), 257-271. doi:10.1037/a0022680

Kazdin, A. E. (2001). Progression of therapy research and clinical application of treatment require better understanding of the change process. *Clinical Psychology: Science and Practice, 8*(2), 143-151. doi:10.1093/clipsy.8.2.143

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*, 1-27. doi:10.1146/annurev.clinpsy.3.022806.091432

Kazdin, A. E. (2008). Evidence-based treatment and practice: New opportunities to bridge

　　clinical research and practice, enhance the knowledge base, and improve patient

　　care. *American Psychologist, 63*(3), 146-159. doi:10.1037/0003-066X.63.3.146

Kendall, P. C., & Hedtke, K. (2006a). *Cognitive-Behavioral Therapy for anxious children:*

　　Therapist Manual. 3rd ed. Ardmore, PA: Workbook Publishing

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

　　Workbook Publishing

Kendall, P. C., Hudson, J. L., Gosch, E., Flannery-Schroeder, E., & 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*(2), 282-297.

　　doi:10.1037/0022-006X.76.2.282

Kendall, P.C., Puliafico, A.C., Barmish, A. J., Choudhury, M. S., Henin, A., & Treadwell, K. S.

　　(2007) Assessing anxiety with the Child Behavior Checklist and the Teacher Report Form.

　　Journal of Anxiety Disorders, *21*(8), 1004-1015. doi:10.1016/j.janxdis.2006.10.012

Langer, D. A., McLeod, B. D., & Weisz, J. R. (2011). Do treatment manuals undermine youth–

　　therapist alliance in community clinical practice? *Journal of Consulting and Clinical*

　　Psychology,79(4), 427-432. doi:10.1037/a0023821

Liber, J. M., McLeod, B. D., Van Widenfelt, B. M., Goedhart, A. W., van der Leeden, Adelinde

　　JM, Utens, E. M., & Treffers, P. D. (2010). Examining the relation between the therapeutic

alliance, treatment adherence, and outcome of cognitive behavioral therapy for children with anxiety disorders. *Behavior Therapy, 41*(2), 172-186. doi:10.1016/j.beth.2009.02.003

Little, J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*(404), 1198-1202. doi:10.2307/2290157

Loeb, K. L., Wilson, G. T., Labovie, E., Pratt, E. M., Hataki, J., Walsh, B. T., & Fairburn, C. G. (2005). Therapeutic alliance and treatment adherence in two interventions for bulimia nervosa: A study of process and outcome. *Journal of Consulting and Clinical Psychology, 73*(6), 1097-1107. doi:10.1037/0022-006X.73.6.1097

Manning, W. H. & DuBols, P. H. Correlational methods in research on human learning. (1962). *Perceptual and Motor Skills 15*, 287-321.

March, J. S., Parker, J. D., Sullivan, K., Stallings, P., & Conners, C. K. (1997). The multidimensional anxiety scale for children (MASC): Factor structure, reliability, and validity. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*(4), 554-565. doi:10.1097/00004583-199704000-00019

March, J. S., & Sullivan, K. (1999). Test–retest reliability of the multidimensional anxiety scale for children. *Journal of Anxiety Disorders, 13*(4), 349–358. doi:10.1016/S0887-6185(99)00009-2

Martin, D. J., Garske, J. P., & Davis, M. K. (2000). Relation of the therapeutic alliance with

outcome and other variables: A meta-analytic review. *Journal of Consulting and Clinical*

Psychology, 68(3), 438-450. doi:10.1037/0022-006X.68.3.438

McLeod, B. D., Islam, N., & Wheat, E. (2013). Designing, conducting, and evaluating therapy

process research. In Comer, S.S. and Kendall, P.C. (Eds.), *The Oxford Handbook of*

Research Strategies for Clinical Psychology (142-164). New York City, NY: 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. doi:10.1037/pas0000037

McLeod, B. D., Southam-Gerow, M. A., Tully, C. B., Rodríguez, A., & Smith, M. M. (2013).

Making a case for treatment integrity as a psychosocial treatment quality indicator for youth

mental health care. *Clinical Psychology: Science and Practice, 20*(1), 14-32.

doi:10.1111/cpsp.12020

McLeod, B. D., & Weisz, J. R. (2004). Using dissertations to examine potential bias in child and

adolescent clinical trials. *Journal of Consulting and Clinical Psychology, 72*(2), 235-251.

doi:10.1037/0022-006X.72.2.235

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. doi:10.1037/0022-006X.73.2.323

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(3), 436-443. doi:10.1080/15374411003691750

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with

 repeated measures and independent-groups designs. *Psychological Methods*, *7*(1), 105-125.

 doi:10.1037//1082-989X.7.1.105

Mowbray, C. T., Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity criteria:

 Development, measurement, and validation. *American Journal of Evaluation*, *24*(3), 315–

 340. doi:10.1177/109821400302400303

Nakamura, B. J., Pestle, S. L., & Chorpita, B. F. (2009). Differential sequencing of cognitive-

 behavioral techniques for reducing child and adolescent anxiety. *Journal of Cognitive*

 Psychotherapy: An International Quarterly. 23(2), 114-135. doi:10.1891/0889-

 8391.23.2.114

Osborne, J.W. (2010). Improving your data transformations: Applying the Box-Cox

 transformation. *Practical Assessment, Research, & Evaluation, 15*(12), 1-9.

Osborne, J. W., & Waters, E. (2002). Four assumptions of multiple regression that researchers

 should always test. *Practical Assessment*, *8*(2), 1-5.

Owen, J., & Hilsenroth, M. J. (2014). Treatment adherence: The importance of therapist

 flexibility in relation to therapy outcomes. *Journal of Counseling Psychology, 61*(2), 280-

 288. doi:10.1037/a0035753

Patel, V., Flisher, A. J., Hetrick, S., & McGorry, P. (2007). Mental health of young people: A global public-health challenge. *The Lancet, 369*(9569), 1302-1313. doi:10.1016/S0140-6736(07)60368-7

Pelham, J., William E, Fabiano, G. A., & Massetti, G. M. (2005). Evidence-based psychosocial treatments for attention-deficit/hyperactivity disorder. *Journal of Clinical Child and Adolescent Psychology, 37*(1), 184-214. doi:10.1080/15374410701818681

Perepletchikova, F. (2009). Treatment integrity and differential treatment effects. *Clinical Psychology: Science and Practice, 16*(3), 379-382. doi:10.1111/j.1468-2850.2009.01177.x

Perepletchikova, F., & Kazdin, A. E. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice, 12*(4), 365-383. doi:10.1093/clipsy.bpi045

Perepletchikova, F., Treat, T. A., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: Analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology, 75*(6), 829-841. doi:10.1037/0022-006X.75.6.829

Piper, W. E., Azim, H. F. A., Joyce, A. S., & McCallum, M. Transference interpretations, therapeutic alliance, and outcome in short-term individual psychotherapy. *Archives of General Psychiatry, 48*(10), 946-953. doi:10.1001/archpsyc.1991.01810340078010

Reynolds, S., Wilson, C., Austin, J., & Hooper, L. (2012). Effects of psychotherapy for anxiety in children and adolescents: A meta-analytic review. *Clinical Psychology Review, 32*(4), 251-262. doi:10.1016/j.cpr.2012.01.005

Rosenthal, D. A., Hoyt, W. T., Ferrin, J. M., Miller, S., & Cohen, N. D. (2006). Advanced

    methods in meta-analytic research applications and implications for rehabilitation

    counseling research. *Rehabilitation Counseling Bulletin, 49*(4), 234-246.

    doi:10.1177/00343552060490040501

Rosnow, R. L., & Rosenthal, R. (1996). Computing contrasts, effect sizes, and counternulls on

    other people's published data: General procedures for research consumers. *Psychological*

    Methods, 1(4), 331-340. doi:10.1037/1082-989X.1.4.331

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the

    art. *Psychological Methods, 7*(2), 147-177. doi:10.1037/1082-989X.7.2.147

Schoenwald, S. K., & Hoagwood, K. (2001). Effectiveness, transportability, and dissemination

    of interventions: What matters when? *Psychiatric Services, 52*(9), 1190-1197.

    doi:10.1176/appi.ps.52.9.1190

Schulte, A. C., Easton, J. E., & Parker, J. (2009). Advances in treatment integrity research:

    Multidisciplinary perspectives on the conceptualization, measurement, and enhancement of

    treatment integrity. *School Psychology Review, 38*(4), 460-475.

Schwab-Stone, M. E., Shaffer, D., Dulcan, M. K., & Jensen, P. S. (1996). Criterion validity of

    the NIMH Diagnostic Interview Schedule for Children Version 2.3 (DISC-2.3). *Journal of*

    the American Academy of Child & Adolescent Psychiatry, 35(7), 878-888.

    doi:10.1097/00004583-199607000-00013

Shaffer, D., Fisher, P., Lucas, C. P., & Dulcan, M. K., & Schwab-Stone, M. E. (1996). The

    NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV):

    Description, differences from previous versions, and reliability of some common diagnoses.

    Journal of the American Academy of Child and Adolescent Psychiatry, 39(1), 28–38.

    doi:10.1097/00004583-200001000-00014

Shrout, P. E., Spitzer, R. L., & Fleiss, J. L. (1987). Quantification of agreement in psychiatric

    diagnosis revisited. *Archives of General Psychiatry, 44*(2), 172-177.

    doi:10.1001/archpsyc.1987.01800140084013

Silverman, W. K., & Albano, A. M. (1996). *The Anxiety Disorders Interview Schedule for DSM-*

    IV-Child and Parent Versions. San Antonio, TX: Graywind Publications.

Silverman, W. K., & Hinshaw, S. P. (2008). The second special issue on evidence-based

    psychosocial treatments for children and adolescents: A 10-year update. *Journal of Clinical*

    Child & Adolescent Psychology, 37(1), 1-7. doi:10.1080/15374410701817725

Silverman, W. K., Pina, A. A., & Viswesvaran, C. (2008). Evidence-based psychosocial

    treatments for phobic and anxiety disorders in children and adolescents. *Journal of Clinical*

    Child & Adolescent Psychology, 37(1), 105-130. doi:10.1080/15374410701817907

Silverman, W.K., Saavedra, L.M., & Pina, A.A. (2001). Test-retest reliability of anxiety

    symptoms and diagnoses with the Anxiety Disorders Interview Schedule for DSM-IV Child

    and Parent versions. *Journal of the American Academy of Child and Adolescent Psychiatry.*

    40(8), 937-944. doi:10.1097/00004583-200108000-00016

Soni, A. (2006). The five most costly children's conditions: Estimates for the US civilian

      noninstitutionalized children, ages 0–17: Statistical brief #242. Rockville, MD, Agency

      for Healthcare Research and Quality, April 2009. Available at

      www.meps.ahrq.gov/mepsweb/data_files/publications/st242/stat242.pdf

Southam-Gerow, M. A., Weisz, J. R., Chu, B. C., McLeod, B. D., Gordis, E. B., & Connor-

      Smith, J. K. (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 & Adolescent Psychiatry, 49(10), 1043-1052. doi:10.1016/j.jaac.2010.06.009

Spielberger, C. D., & Edwards, C. D. (1973). *STAIC preliminary manual for the state-trait*

      anxiety inventory for children ("how I feel questionnaire") Consulting Psychologists

      Press.

Tabachnick, B. C., & Fidell, L. S. (2007). *Using multivariate statistics* (5th edition). Boston, MA:

      Allyn and Bacon.

Tucker, L.R., Damarin, F., & Messick, S. (1966). A base-free measure of change.

      Psychometrika, 31(4), 457-473. doi:10.1007/BF02289517

US Department of Health and Human Services, National Institute of Health, National Institute of

      Mental Health (2011). Use of mental health services and treatment among children.

      Retrieved from the National Institute of Mental Health website:

      http://www.nimh.nih.gov/statistics/1NHANES.shtml

Waltz, J., Addis, M. E., Koerner, K., & Jacobson, N. S. (1993). Testing the integrity of a

      psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting*

      and Clinical Psychology, 61(4), 620-630. doi:10.1037/0022-006X.61.4.620

Webb, C. A., DeRubeis, R. J., & Barber, J. P. (2010). Therapist adherence/competence and

    treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical*

    Psychology,78(2), 200-211. doi:10.1037/a0018912

Webb, C. A., DeRubeis, R. J., Dimidjian, S., Hollon, S. D., Amsterdam, J. D., & Shelton, R. C.

    (2012). Predictors of patient cognitive therapy skills and symptom change in two

    randomized clinical trials: The role of therapist adherence and the therapeutic

    alliance. *Journal of Consulting and Clinical Psychology, 80*(3), 373-381.

    doi:10.1037/a0027663

Weersing, V. R., & Weisz, J. R. (2002). Mechanisms of action in youth psychotherapy. *Journal*

    of Child Psychology and Psychiatry, 43(1), 3-29. doi:10.1111/1469-7610.00002

Weisz, J. R., Doss, A. J., & Hawley, K. M. (2005). Youth psychotherapy outcome research: A

    review and critique of the evidence base. *Annual Review of Psychology, 56*, 337-363.

    doi:10.1146/annurev.psych.55.090902.141449

Weisz, J. R., & Gray, J. S. (2008). Evidence-based psychotherapy for children and adolescents:

    Data from the present and a model for the future. *Child and Adolescent Mental*

    Health, 13(2), 54-65. doi:10.1111/j.1475-3588.2007.00475.x

Weisz, J. R., Jensen-Doss, A., & Hawley, K. M. (2006). Evidence-based youth psychotherapies

    versus usual clinical care: A meta-analysis of direct comparisons. *American*

    Psychologist,61(7), 671-689. doi:10.1037/0003-066X.61.7.671

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. doi:10.1037/a0013877

Weisz, J. R., Ugueto, A. M., Cheron, D. M., & Herren, J. (2013). Evidence-based youth psychotherapy in the mental health ecosystem. *Journal of Clinical Child & Adolescent Psychology, 42*(2), 274-286. doi:10.1080/15374416.2013.764824

White, K., Kendrick, T., & Yardley, L. (2009). Change in self-esteem, self-efficacy and the mood dimension of depression as potential mediators of the physical activity and depression relationship: Exploring the temporal relation of change. *Mental Health and Physical Activity, 2*(1), 44-52. doi:10.1016/j.mhpa.2009.03.001

Williams, R. H., Zimmerman, D. W., & Mazzagatti, R. D. (1987). Large sample estimates of the reliability of simple, residualized, and base-free gain scores. *Journal of Experimental Education, 55*(2), 116-118. doi:10.1080/00220973.1987.10806443

Wood, J. J., Piacentini, J. C., Bergman, R. L., McCracken, J., & Barrios, V. (2002). Concurrent validity of the anxiety disorders section of the anxiety disorders interview schedule for DSM-IV: Child and parent versions. *Journal of Clinical Child and Adolescent Psychology, 31*(3), 335-342. doi:10.1207/S15374424JCCP3103_05

Wood, J. J., Piacentini, J. C., Southam-Gerow, M., Chu, B. C., & Sigman, M. (2006). Family cognitive behavioral therapy for child anxiety disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 45*(3), 314-321. doi:10.1097/01.chi.0000196425.88341.b0

Vita

Ruben Martinez was born on April 21, 1989 in Muncie, Indiana. He graduated from Greenwood Community High School, Greenwood, Indiana in 2007. He then graduated from Indiana University in May 2011 with a Bachelor of Arts in Psychology. Subsequently, he worked for two years as a research assistant at the Cognitive Development Lab and a lab/project coordinator in the Training Research and Implementation in Psychology labs at Indiana University before entering the Clinical Psychology doctoral program at Virginia Commonwealth University in Richmond, Virginia.