Virginia Commonwealth University
VCU Scholars Compass

Theses and Dissertations | Graduate School

2019

Discriminating between Technical and Global Competence in CBT Programs for Youth Anxiety

Jennifer Cecilione
Virginia Commonwealth University

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

 Part of the Clinical Psychology Commons

© The Author

Downloaded from
https://scholarscompass.vcu.edu/etd/6130

This Thesis is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.

Discriminating between Technical and Global Competence in CBT Programs for Youth Anxiety

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

By: Jennifer Cecilione

Bachelor of Arts, University of Richmond, 2015

Director: Bryce D. McLeod, Ph.D.

Associate Professor

Department of Psychology

Virginia Commonwealth University

Richmond, Virginia

December, 2019

Acknowledgments

I would like to thank Dr. Bryce D. McLeod for his mentorship, guidance, and feedback throughout this process, Dr. Roxann Roberson-Nay and Dr. Michael Southam-Gerow for kindly agreeing to serve on my committee and Drs. Meghan Smith, Adriana Rodriguez, Carrie Tully, and Nadia Islam for their dedication to hours of coding.

Table of Contents

Abstract

DISCRIMINATING BETWEEN TECHNICAL AND GLOBAL COMPETENCE IN CBT PROGRAMS FOR YOUTH ANXIETY

By: Jennifer Cecilione, 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, 2019.

Major Director: Bryce D. McLeod, Ph.D.
Professor
Department of Psychology

Therapist competence refers to the skillfulness and responsiveness demonstrated when delivering an intervention and is an important factor to consider in the training and evaluation of therapists. However, competence research is sparse, especially in the youth psychosocial treatment field. A primary discrepancy is whether technical (related to interventions associated with a specific treatment program) and global (general clinical expertise) competence can be measured as distinct dimensions of competence. The goal of the current study was to determine whether instruments of technical (Cognitive-Behavioral Treatment for Anxiety in Youth Competence Scale; CBAY-C; McLeod et al., 2018) and global (Global Therapist Competence Scale for Youth Psychosocial Treatment; G-COMP; Brown et al., 2018) competence assessed distinct constructs in the context of cognitive-behavioral treatment (CBT) for youth anxiety. Treatment sessions (n = 359 SMT; n = 244 MMT) from 38 youth participants (n = 16 MMT; n = 22 SMT; M age = 9.84 years, SD = 1.65; 47.4% female, 60.5% Caucasian) from an effectiveness study were coded by two coders in two CBT programs (modular manualized treatment [MMT] and standard manualized treatment [SMT]) for youth anxiety using observational coding systems designed to assess competence, adherence, and alliance. The average intraclass correlations

(ICC[2,2]) for CBAY-C model items was .53 (*SD* = .23) in the MMT condition and .71 (*SD* = .08) in the SMT condition. The average ICC(2,2) for G-COMP items was .53 (*SD* = .04) in the MMT condition and .69 (*SD* = .05) in the SMT condition. The average correlation between technical (CBAY-C) and global competence (G-COMP) subscale scores was *r* = .59 in the MMT condition and *r* = .73 in the SMT condition. The findings suggested there was insufficient evidence to demonstrate that technical competence could be consistently measured as distinct from all dimensions of global competence across two different CBT programs. Future research should examine the potential distinction between technical and global competence in treatment programs other than CBT.

Introduction

Within the past couple of decades, there has been an increased focus placed upon competence-based practice in professional psychology (Fouad et al., 2009; Humphreys, Crino, & Wilson, 2018; Kaslow, 2004). For instance, the American Psychological Association (APA) code of ethics specifically highlights practicing in a "competent" manner as an ethical standard of all psychologists (2002). To meet this goal, it follows that therapist competence (hereafter referred to as *competence*) should be considered in the training and evaluation of therapists. Yet, competence is understudied, especially in the youth psychosocial treatment field. It is important to study the competence of therapists who see child clients separately from competence in the adult treatment field, since working with children may require the consideration of unique factors (e.g., delivering treatment in a developmentally appropriate manner; Eyberg, Schuhmann, & Rey, 1998).

Briefly, competence is a combination of both the skillfulness and responsiveness a therapist demonstrates when delivering an intervention (e.g., Barber et al., 2007; Hogue et al., 2008). More specifically, two primary dimensions of competence have been proposed—technical and global competence (e.g., Barber et al., 2007). Technical competence refers to a therapist's skillfulness and responsiveness in the context of implementing a specific intervention found in a treatment program (Barber et al., 2007; Hogue et al., 2008; McLeod et al., 2018). Technical competence might be more useful to consider when evaluating and training therapists in specific specialties (e.g., cognitive-behavioral therapy or CBT; Sharpless & Barber, 2009). Conversely, global competence refers to a therapist's general clinical skills and judgement that cut across various treatment modalities, interventions, and programs (Barber et al., 2007; Sharpless & Barber, 2009). Global competence might be more useful to consider when therapists are

evaluated for internship and/or independent practice readiness (Sharpless & Barber, 2009). Since there might be certain training and/or evaluation situations in which one dimension of competence is more appropriate to consider over another, it is important to understand the two as distinct dimensions. However, there are several inconsistencies and gaps in the current literature that obfuscate the understanding of these two competence dimensions.

One major barrier to studying technical and global competence is the lack of a consensus amongst researchers regarding how to define these two competence dimensions. For example, some researchers have conceptualized technical competence more broadly and postulated that it encompasses the processes and structures involved in implementing treatment goals (e.g., "partnership working", "right developmental level"; Stallard et al., 2014). Conversely, other researchers have defined technical competence as the skillfulness and responsiveness demonstrated in the context of implementing discrete practice elements (i.e., distinct clinical techniques utilized as components of a larger intervention program; Chorpita & Daleiden, 2009). Those who have defined technical competence more broadly tend to have items in their instruments that incorporate several elements of treatment (e.g., assigning homework, reviewing homework, agenda setting, etc.) into one competence item (e.g., "CBT structure"; Bjaastad et al., 2016). Alternatively, those who have defined technical competence more specifically and narrowly have items in their instruments that correspond to the implementation of practice elements unique to a certain treatment program (e.g., "coping plan", "problem-solving"; McLeod et al., 2018) and the specific interventions used to deliver them (e.g., "collaborative teaching"; McLeod et al., 2018).

Similarly, some researchers have conceptualized global competence broadly as a therapist's general or "overall" competence demonstrated throughout a session via characteristics

such as empathy (Hogue et al., 2008) and interpersonal skills (e.g., Gutermann et al., 2015). When global competence is defined more broadly, it is often measured with one item (e.g., "overall session competence"; Gutermann et al., 2015) or a few items (e.g., "skill", "responsiveness" and "overall competence"; Hogue et al., 2008). Only one instrument, the Global Therapist Competence Scale for Youth Psychosocial Treatment (G-COMP; Brown et al., 2018), has been created for the sole purpose of measuring global competence in youth treatment. Brown et al. (2018) described five theoretically driven dimensions that comprise global competence (i.e., alliance building, positive expectancies, focusing treatment, instigating change, and responsiveness). Each of these five domains of global competence serve as the level-one items of the G-COMP (hereafter referred to as "G-COMP items" or "G-COMP scores").

Since a primary aim of the current study was to determine if technical and global competence can be distinguished from one another, it was important to use instruments that mapped onto clear definitions of these constructs. One instrument of technical competence, the CBT for Anxiety in Youth Competence Scale (CBAY-C; McLeod et al., 2018) contains "model" items (e.g., "exposure debrief", "fear ladder") that assess competence demonstrated when implementing each practice element. Because the CBAY-C operationalized technical competence precisely and provided specific potential targets for therapist evaluation and training (i.e., practice elements), CBAY-C model items were used in the current study's analyses as a measure of technical competence. Additionally, the G-COMP items represent a specific operationalization of global competence that is conceptually distinct from technical competence. Thus, G-COMP items were used in the current study's analyses as a measure of global competence.

Besides the heterogeneity of definitions and measurement of technical and global competence, few studies have extensively evaluated the psychometric properties of competence scores that are used in youth treatment. Before conclusions can be drawn about the distinction between competence dimensions, there are several key psychometric properties that should be considered. First, because many competence instruments (including the instruments considered in the present study) are observer-reported, it is important to evaluate inter-rater reliability. Inter-rater reliability is the extent to which observers agree on the scores they assign when coding participants' performance on an instrument (Kazdin, 2016) and is often measured via intra-class correlations (ICC[2,2]s; Koo & Li, 2016). If inter-rater reliability is low, then competence scores generated from that instrument would not be considered reliable estimates of the construct. Low inter-rater reliability might mean that a significant proportion of the variance in competence scores could be due to coder differences, which would limit conclusions that could be made regarding competence in that sample.

The inter-rater reliability of competence scores has been previously assessed in the context of youth CBT programs. For instance, Bjaastad et al. (2016) found a mean ICC(2,2) for all competence items of .53 (*SD* = .10). Hogue et al. (2008) found a mean ICC(2,2) of .43 (*SD* = .19) of all the competence items in a CBT condition. Moreover, the mean ICC(2,2) for all CBAY-C scores was ICC(2,2) = .67 (*SD* = .11; McLeod et al., 2018), while the mean ICC(2,2) for G-COMP scores was ICC(2,2) = .70 (*SD* = .07; Brown et al., 2018). These ICCs suggest that items on the CBAY-C and G-COMP can be reliably assessed using two coders. Overall, these findings suggest that the inter-rater reliability of competence items in the literature ranges from "fair" to "good" (Cicchetti, 1994).

Moreover, construct validity is also important to consider when evaluating scores of technical and global competence instruments. Construct validity is the degree to which scores on an instrument represent the construct they are purported to assess (Kazdin, 2016). If scores on an instrument are not actually measuring technical or global competence, then using that instrument to evaluate technical or global competence would not generate meaningfully interpretable results. Construct validity cannot be determined by only one correlation between scores on two instruments. Instead, construct validity is established by examining patterns of associations with scores on similar instruments, which often involves assessing different types of validity (Kazdin, 2016). The current study considered convergent and discriminant validity as key psychometric properties that could support the construct validity of CBAY-C and G-COMP scores. Convergent validity refers to the degree to which scores on instruments that measure similar constructs are associated with one another. Conversely, discriminant validity refers to the extent to which instruments designed to measure distinct constructs produce scores that are not associated with each other (Kazdin, 2016).

A few studies have examined the association between technical and global competence scores. For instance, Brown et al. (2018) found that G-COMP scores were associated with the CBAY-C total score: Alliance Building (r = .26), Positive Expectancies (r = .24), Focusing Treatment (r = .53), Instigating Change (r = .52), and Responsiveness (r = .26) (M = .36, SD = .15). Hogue et al. (2008) also compared scores on an "overall competence" item to the average score of technical competence items in CBT sessions (r = .68) and Multi-Dimensional Family Therapy (MDFT) sessions (r = .79). Additionally, Bjaastad et al. (2016) found an association between a global rating of competence and the sum score of all four competence items (r = .88). McLeod et al. (2018) also found that CBAY-C model item scores were correlated with overall

therapist skillfulness scores (*r's* = .59 - .91, *M* = .77, *SD* = .12) and overall therapist responsiveness scores (*r's* = .73 - .89; *M* =.82, *SD* = .08). Overall, these correlations amongst competence scores can be classified as "large" (Rosenthal and Rosnow, 1984), while most correlations were at a level that suggested conceptual overlap (*r*'s > .7; Kline, 1979). Brown et al. (2018)'s findings were the only findings that did not suggest that technical and global competence scores correlated at a level of conceptual overlap. These mixed results do not provide a clear answer as to whether scores on technical and global competence items can discriminate technical from global competence.

Competence scores have also been compared to adherence (i.e., the extent to which a therapist delivers an intervention as it was originally intended to be implemented; Barber et al., 2007) scores to assess discriminant validity of competence scores. For example, the mean inter-item correlation between CBAY-C scores and scores on corresponding items of the Cognitive Behavioral Therapy Adherence Scale for Youth Anxiety (CBAY-A; Southam-Gerow et al., 2016) was *r* = .43 (*SD* = .20; McLeod et al., 2018). Similarly, the mean inter-item correlation between G-COMP scores and CBAY-A total scores was *r* = .39 (*SD* = .18; Brown et al., 2018). Hogue et al. (2008) also found that the average score of adherence items was correlated with the average score of technical competence items in the CBT condition (*r* = .42) and in the MDFT (*r* = .17) condition. Additionally, the "overall competence" score was also correlated with the average score of adherence items in CBT (*r* = .50) and MDFT (*r* = .23) conditions (Hogue et al., 2008). Correlations between scores on instruments of competence and adherence were found in other studies as well (*r* = .79, Bjaastad et al., 2016; *r* = .65; Gutermann et al., 2015). These correlations suggest that competence and adherence are strongly related constructs. Overall, correlations between competence and adherence scores were smaller in magnitude than

correlations between technical and global competence scores, which supports the discriminant validity of these competence scores. Yet, there was one instance in which competence scores were more strongly associated with adherence scores than other competence scores (i.e., Brown et al., 2018); thus, further examination of competence scores' validity is needed.

Previous studies have also compared competence scores to scores on instruments of alliance (i.e., the quality of the relationship between the therapist and client; McLeod & Weisz, 2005) to examine the discriminant validity of competence scores. For instance, the mean inter-item correlation between CBAY-C and Therapy Process Observational Coding System-Alliance Scale (TPOCS-A; McLeod & Weisz, 2005) total scores was r = .21 (SD = .14; McLeod et al., 2018), and the mean inter-item correlation between G-COMP and TPOCS-A scores was r = .32 (SD = .11 Brown et al., 2018). Additionally, Brown et al. (2018) found a mean inter-item correlation of r = .22 (SD = .04) between scores on the last observation of G-COMP items and scores on the Therapeutic Alliance Scale for Children (TASC), which is an end-of-treatment youth-reported alliance instrument (Shirk & Saiz, 1992). Hogue et al. (2008) also found that Vanderbilt Therapeutic Alliance Scale-Revised (VTAS-R; Hogue et al., 2006) scores were correlated with technical competence scores in the MDFT condition (r = .40), "overall competence" scores in the CBT condition (r = .31), and "overall competence" scores the MDFT condition (r = .36). VTAS-R scores were not significantly correlated with technical competence scores in the CBT condition (Hogue et al., 2008). Overall, the correlations between competence and alliance scores were smaller in magnitude than correlations between technical and global competence scores, which supports the discriminant validity of these competence scores. Moreover, that CBAY-C and G-COMP scores were more closely associated with adherence

scores than they were with alliance scores supports the construct validity of these competence scores.

In sum, there is insufficient evidence to conclude that technical and global competence can be reliably and validly measured as distinct constructs. Thus, the current study's aim was to determine whether technical and global competence could be distinguished from one another in a psychometrically sound manner. In the present study, *technical competence* was defined as a therapist's skillfulness and responsiveness when implementing a specific intervention technique (i.e., practice element) found in a particular treatment program (Barber et al., 2007; Hogue et al., 2008; McLeod et al., 2018). Additionally, *global competence* was defined as a therapists' general "clinical acumen" that permeates their interventions (Barber et al., 2007; Sharpless & Barber, 2009) and cuts across various treatment modalities, interventions and programs (Brown et al., 2018). The current study utilized CBAY-C model items as a measure of technical competence and G-COMP items as a measure of global competence, as these items best mapped on to the definitions of each construct.

The present study assessed several psychometric properties of both the CBAY-C and G-COMP as utilized to measure technical and global competence of therapists in a randomized effectiveness trial (Child STEPS; Weisz et al., 2012). Child STEPs participants were children aged 7-13 with presenting problems of either anxiety, depression or conduct disorder receiving either modular manualized treatment (MMT), standard manualized treatment (SMT), or usual care. Only cases of anxiety were used in the present study, as the focus was evaluating competence in the context of CBT for youth anxiety. Additionally, only the MMT and SMT conditions were included in analyses. Because the usual care condition was not a specific treatment program, it would be inappropriate to assess therapists' technical competence of

implementing CBT in the usual care condition. Additionally, both the CBAY-C and G-COMP were used to code competence demonstrated in Child STEPs sessions. Because these instruments' operational definitions best mapped on to the provided definitions of technical and global competence, they allowed for meaningful conclusions about technical and global competence to be drawn from the current study's results. Lastly, since the CBAY-C was created specifically to measure technical competence in the context of CBT for youth anxiety, it was especially appropriate to use with the current sample.

To address the primary aim of the current study (i.e., to determine if technical and global competence could be measured distinctly in CBT for youth anxiety), the following analyses were conducted. First, the inter-rater reliability of the CBAY-C model items as well as the G-COMP items were examined to determine if these instruments could be used by two coders to arrive at similar scores. It was expected that inter-rater reliability of competence scores would be at least "good" ($ICC(2,2) \geq .60$; Cicchetti, 1994) in both the MMT and SMT conditions.

Then, competence (CBAY-C and G-COMP) scores were correlated with one another, and it was expected that the correlations would be "large" (r's $\geq .36$; Rosenthal & Rosnow, 1984) but not redundant (Kline, 1979) in both study conditions. Next, patterns of correlations between CBAY-C model scores and scores on instruments of related constructs (i.e., adherence and alliance) were assessed. Similarly, G-COMP scores were also compared to scores on instruments of related constructs (i.e., adherence and alliance) as well as to CBAY-C model scores. It was expected that competence (CBAY-C and G-COMP) scores would be most strongly correlated with other competence (CBAY-C and G-COMP) scores. It was also expected that competence (CBAY-C and G-COMP) scores would be strongly correlated with adherence (CBAY-A) scores and less strongly correlated with alliance (TPOCS-A) scores. This pattern of associations has

been established in previous studies (e.g., Brown et al., 2018; Hogue et al., 2008; McLeod et al., 2018).

Literature Review

The American Psychological Association (APA) code of ethics emphasizes the importance of practicing psychology in a competent manner (2002). Increasingly, mental health services are becoming focused on accountability. Stakeholders with a vested interest in the outcome produced by psychosocial treatments (hereafter referred to as *treatments*) are increasingly demanding that the delivery of treatments meet certain criteria related to the quality of services provided to consumers. This focus on quality is due partly to the increased focus on accountability and the competence-based movement in psychology (Fouad et al., 2009; Humphreys et al., 2018; Kaslow, 2004). An important goal of this movement is to improve the quality of mental health services by providing guidelines for training and practice related to the delivery of treatments. For instance, therapists must demonstrate a certain level of competence before they are deemed ready for internship and/or independent practice. Similarly, whether a graduate psychology program earns APA accreditation is based in part on whether it can produce "competent" graduates (Kaslow, 2004). Clients of therapists should also be concerned about their therapist's level of competence, so they can ensure that they are spending their resources on quality care from competent providers (Sharpless & Barber, 2009). Thus, therapist competence (hereafter referred to as *competence*) is important for helping to ensure quality of mental health care.

Definition of Competence

Competence research is sparse in the youth treatment field. It important to examine competence in the context of youth treatment separately from adult treatment, because there are

several factors that should be considered in the delivery of youth treatment that are distinct from adult treatment (e.g., developmental stage; Eyberg et al., 1998; self-referral versus parent-referral; McLeod & Weisz, 2005; Podell et al., 2013) that could affect what constitutes competence in each setting. Thus, the current study focused only on competence in the context of youth treatment.

Though competence represents an important concept, the term has not always been consistently defined. Still, youth treatment researchers have typically agreed on two broad characteristics of competence: *skillfulness* (e.g., Bjaastad et al., 2016; Brown et al., 2018; Gutermann et al., 2015; Hogue et al., 2008; Marvin et al., 2016; McLeod et al., 2018) and *responsiveness* (e.g., Brown et al., 2018; Hogue et al., 2008; McLeod et al., 2018; Resko et al., 2012). Example elements of competence in the youth treatment field include: the appropriateness and timing of interventions (Hogue et al., 2008; McLeod et al., 2018), efforts to manage the therapeutic relationship and/or encourage change (Brown et al., 2018), consideration of variables that might be relevant to the therapeutic context (e.g., client age, symptom severity, stage of treatment; Gutermann et al., 2015), and aspects related to communication and technical abilities (Resko et al., 2012). Generally, competence can be thought of as the quality of treatment delivery that encompasses the amount of skill and judgement a therapist demonstrates when implementing an intervention (Perepletchikova & Kazdin, 2005).

Some researchers have also posited a theoretical distinction between two dimensions of competence. Technical competence (also known as “limited domain” or “model-specific” competence; Barber et al., 2007) may be more relevant to specific treatment programs (e.g., conducting exposures skillfully is especially relevant to CBT) and be particularly useful when evaluating competence in the context of a certain specialty (Sharpless & Barber, 2009).

Conversely, global competence (also known as “common-factors”, Castonguay, 1993; “foundational”, or “general competence”, Spruill et al., 2004) may be more general and applicable to all treatments (e.g., the ability to build rapport with clients) and seems to be what therapists are judged on when they are being evaluated for internship or independent practice readiness (Sharpless & Barber, 2009). Distinguishing between technical and global competence is important because it could have implications for the evaluation and training of future therapists (Brown et al., 2018; Sburlati et al., 2011).

Technical competence describes a therapist’s skillfulness and responsiveness when implementing specific interventions found in a particular treatment program and has been the primary focus of competence research thus far (Barber et al., 2007; Hogue et al., 2008; McLeod et al., 2018). However, a discrepancy in the literature exists regarding how to define technical competence. Some researchers have defined technical competence broadly by “standard items” that describe molar treatment goals (e.g., “CBT structure”; Bjaastad et al., 2016 and “family interaction interventions”; Hogue et al., 2008) and processes by which those goals are achieved (e.g., “implementation of techniques”; Gutermann et al., 2015 and “enjoyable and engaging”; Stallard et al., 2014). Whereas, other researchers have defined technical competence as the skillfulness and responsiveness demonstrated in the implementation of specific, core practice elements (i.e., discrete clinical techniques used as part of a larger intervention plan; Chorpita & Daleiden, 2009) that are key to a particular treatment program (e.g., “model” items, such as “problem solving”; McLeod et al., 2018).

To distinguish between technical and global competence, the ways in which they are defined should not conflate the two constructs. Conceptualizing technical competence with a broad definition may unintentionally capture aspects of global competencies (e.g., rapport and

alliance building); therefore, it seems more precise to define technical competence in terms of competencies directly related to discrete practice elements. Thus, the current study considered technical competence as the skillfulness and responsiveness demonstrated in the implementation of specific practice elements found in treatment programs. Examples of technical competence defined in this way include the skillfulness and responsiveness demonstrated when conducting an exposure with a client or when helping a client challenge their negative thoughts (McLeod et al., 2018).

Alternatively, global competence refers to a therapist's general ability to make clinical judgements that permeates their interventions (Barber et al., 2007; Sharpless & Barber, 2009) and cuts across various treatment modalities, interventions, and programs (Brown et al., 2018). Hence, global competence is not skillfulness or responsiveness specifically associated with a discrete practice element of a treatment program (e.g., how skillfully a therapist teaches a child how to problem-solve). Yet, another discrepancy regarding how to define global competence exists in the literature. Some researchers have described global competence more broadly than others. For instance, Gutermann et al. (2015) defined global competencies as those that demonstrate the therapist's overall capacity to support clients by properly structuring sessions or by demonstrating advanced interpersonal skills that are separate from competencies associated with a specific treatment. Similarly, Hogue et al. (2008) described global competence as a method for measuring competence that involves the therapist's general competence as displayed in a session via skill, empathy and nonverbal behaviors.

Alternatively, Brown et al. (2018) defined global competence as therapists' ability (across various treatment modalities) to manage the therapeutic relationship, encourage change in the client, and implement intervention(s) at appropriate times for a given client. More

specifically, they posited that global competence is composed of five distinct, yet related domains. First, “alliance building” is comprised of elements such as empathy and demonstrating understanding (Norcross, 2002, 2011), therapeutic sincerity (Orlinsky & Howard, 1987), as well as caring, warmth, and acceptance of the client (Castonguay & Beutler, 2006). The second domain of global competence, “positive expectancies”, is comprised of elements such as bolstering the client's beliefs about the helping process (Frank, 1971). Third, “focusing treatment” is characterized by elements such as structuring a focused treatment session (Castonguay & Beutler, 2006). Fourth, “instigating change” is comprised of elements like the encouragement of emotional reactions (Frank, 1971), and guided self-exploration (Castonguay & Beutler, 2006). Lastly, “responsiveness” is thought of as handling resistance, tailoring treatment (Norcross, 2002, 2011), and demonstrating appropriate responsiveness throughout treatment (Castonguay & Beutler, 2006).

A potential problem with conceptualizing global competence too generally with one or two items is that it might increase the likelihood of conflating technical and global competence. For instance, considering the general ability of a therapist could be conceptualized as a summary of all competencies (including technical competencies). Brown et al. (2018)’s method of defining global competence seems preferable; the five proposed domains are applicable to all therapeutic interactions and are clearly distinguished from competencies associated with discrete practice elements. Hence, the current study considered global competence as the skillfulness and responsiveness that is integral to all therapeutic interactions, as represented by five primary domains (i.e., alliance building, positive expectancies, focusing treatment, instigating change, and responsiveness).

In sum, definitions of technical and global competence should clearly differentiate between the two dimensions of competence. If definitions of technical and global competence overlap in content, it follows that the measurement of the two dimensions will overlap in content as well. If the instruments designed to measure technical and global competence contain overlapping content, it seems impossible to differentiate the two dimensions in the training and evaluation of therapists. Hence, the clarity and precision of definitions are important factors to consider when assessing the distinction between technical and global competence.

Measurement of Technical and Global Competence

Technical and global competence represent important dimensions of competence. However, the conceptualization and measurement of technical and global competence has been inconsistent in the youth treatment field. If technical and global competence are to be understood and utilized in a meaningful way in research and practice, they need to be measured as distinct constructs. Therefore, it is important to determine how instruments have attempted to assess these constructs.

Operational Definitions of Technical Competence

While technical competence is broadly considered as skillfulness and responsiveness related to particular treatment programs, operational definitions of technical competence in existing competence instruments are varied. Operational definitions of technical competence seem to fall into one of two categories: measuring technical competence more broadly (i.e., treatment goals and process items) or more narrowly (i.e., items related to discrete practice elements that are unique to certain treatment programs).

Some instruments that operationalize technical competence more broadly contain items that correspond to molar treatment goals, which are broad over-arching, and integrative goals

that are comprised of "integrated intervention techniques" that cut across multiple sessions (Hogue et al., 2008, p. 138). For example, the Therapist Behavior Rating Scale-Competence (TBRS-C; Hogue et al., 2008) is an observer-rated instrument that is used to assess competence in individual cognitive-behavioral therapy (ICBT) and multidimensional family therapy (MDFT) for adolescent substance abuse. The TBRS-C measured technical competence with items that correspond to molar therapeutic goals of either ICBT (i.e., "establishing a working relationship", "drug-use monitoring", "behavioral skills training", "cognitive therapy techniques", and "increasing prosocial behavior") or MDFT (e.g., "adolescent interventions", "parent interventions", "family interaction interventions", and "extrafamilial interventions"; Hogue et al., 2008, pgs. 140-141). The TBRS-C items were not competence ratings of individual practices (e.g., exposure, psychoeducation); rather, they were competence ratings related to multiple practices combined.

Another instrument that assessed technical competence more broadly is the CBT Scale for Children and Young People (CBTS-CYP; Stallard et al., 2014), which was intended to assess competence in CBT for youth with internalizing psychopathology. The CBTS-CYP included items that evaluate competence in the implementation of methods and in the utilization of process of using CBT with youth. The CBTS-CYP "method" items (i.e., "assessment and goals", "behavioral techniques", "cognitive techniques", "discovery", "emotional", "formulation", and "general skills") appear to capture technical competence as it relates to molar treatment goals. For example, "behavioral techniques" describes instances when the therapist uses a variety of practices to encourage understanding and therapeutic change (Stallard et al., 2014, p. 274). Whereas, the "process" items (i.e., "partnership working", "right developmental level", "empathy", "creative", "investigation", "self-efficacy", "enjoyable and engaging") are purported

to assess elements of processes involved in implementing CBT with youth (Stallard et al., 2014, p. 273).

Moreover, the Competence and Adherence Scale for CBT (CAS-CBT; Bjaastad et al., 2016) was also designed to measure technical competence in CBT sessions for youth anxiety by assessing competencies of treatment goals and CBT processes. CAS-CBT items were categorized into two subscales: "CBT structure and session goals" (e.g., "cognitive therapy structure") and "process and relational skills" (e.g., "flexibility"; Bjaastad et al., 2016, p. 6). The "CBT structure and session goals" items do not isolate competencies associated with discrete practices or skills, but instead combine several skills (e.g., "homework review and planning new homework", "structure and progress", "parental involvement") into single competence scores. Similarly, the "process and relational skills" items do not capture competencies of discrete practices, but instead capture competencies associated with processes involved in implementing CBT (e.g., "collaboration", "positive reinforcement").

The Global Rating of Motivational Interviewing Therapist (GROMIT; Resko et al., 2012) has also been used to measure competence as it relates to treatment goals and processes in a brief motivational interviewing and skills training intervention for alcohol misuse and violent behaviors. The GROMIT was comprised of two factors: "empathic counseling style representative of the tenants of motivational interviewing" (e.g., "understanding the client's point of view", "expressed approval of the client") and "empowerment and the therapist's skillfulness in negotiating power issues" (e.g., "therapist did not assume the expert role", "therapist seemed genuine"; Resko et al., 2012, p. 7). Though not clearly labeled as such, some items seemed to capture competence related to molar treatment goals (e.g., "guided the client toward change

talk"), while others seemed to assess competence related to the processes involved when striving towards the treatment goals (e.g., "did not steam roll the client").

Another instrument, the Therapeutic Competence Scale (TCS; Gutermann et al., 2015), has also been used to measure technical competence more broadly in a Developmentally Adapted Cognitive Processing Therapy (D-CPT) for post-traumatic stress disorder (PTSD). The TCS included items adapted from the Cognitive Therapy Scale (CTS; Weck et al., 2011) to better fit PTSD in adolescents (e.g., "clarity of communication", "reviewing homework") as well as items specific to D-CPT (e.g., "dealing with severe stress", "facilitating cooperation"). Some of the TCS items appear to capture competencies of molar treatment goals (e.g., "focus on the cognitive model"), while others seem to assess competence of the processes that would be necessary to reach treatment goals (e.g., "interpersonal effectiveness").

Operationalizing technical competence with broad items might make it difficult to isolate specific competencies related to a particular treatment program, as multiple practices may be utilized in achieving a treatment goal (e.g., a therapist might utilize psychoeducation and engagement practices to "establish a working relationship"; Hogue et al., 2008). Hence, by rating technical competence on the level of treatment goals, one might not capture competencies associated with discrete practice elements of a treatment program, which might be particularly useful in therapist training and evaluation. Similarly, it seems that items assessing the competence of processes involved in reaching treatment goals might not isolate technical competencies unique to a treatment program. For instance, process items such as "empathy" (CBTS-CYP; Stallard et al., 2014), "seems genuine" (GROMIT; Resko et al., 2012), "clarity of communication" (TSC; Gutermann et al., 2015) and "flexibility" (CAS-CBT; Bjaastad et al., 2016) might be capturing global competencies related to all youth treatments (e.g., alliance

building). Overall, it seems that measuring technical competence with broad treatment goal and process items is not the most precise way to measure technical competence and is potentially susceptible to the conflation of the two dimensions of competence.

Another way that researchers have conceptualized technical competence is to assess it in the context of discrete practice elements. One instrument that was built around this conceptualization of technical competence is the CBT for Anxiety in Youth Competence Scale (CBAY-C; McLeod et al., 2018). The CBAY-C items are categorized into 4 groups: "standard interventions" (i.e., interventions common to CBT), "model interventions" (i.e., theory-driven, fundamental interventions specific to ICBT for youth anxiety), "delivery" (i.e., the manner by which model interventions were implemented or delivered), and "global" items (i.e., "skillfulness" and "responsiveness").

The CBAY-C model items were designed to measure technical competence of discrete practice elements in youth CBT; they provide a clear picture of a therapists' specific competencies in implementing core practice elements of a treatment program and allow for the analysis of technical competence separately from global competence (McLeod et al., 2018). Operationalizing technical competence by interventions and processes is common to most technical competence instruments used in youth treatment (e.g., Bjaastad et al., 2016; Hogue et al., 2008; Stallard et al., 2014). Yet, the CBAY-C model items offer a greater degree of definitional specificity than previous measures of technical competence. These aspects of measuring technical competence with the CBAY-C model items map onto the larger goal of examining technical competence in a way that can inform the evaluation of therapists by providing specific targets for improvement.

Operational Definitions of Global Competence

Competence research in the youth treatment field is especially sparse with regards to global competence. Conceptualizations of how global competence should be operationalized seem to fall into one of two categories: measuring global competence with one or a few global items or with multiple items that capture specific domains of global competence.

Some researchers assess global competence by appending one or a few items to the end of a technical competence instrument. For example, the TCS (Gutermann et al., 2015) included one item intended to assess global competence (i.e., "overall session competence"). Similarly, the TBRS-C included three items that measure competence via the "global rating method" (i.e., "skill", "responsiveness" and "overall competence"; Hogue et al., 2008). However, measuring global competence with a few, brief items might not clearly discriminate between technical and global competence. Based on information provided in these manuscripts, it is unclear what criteria are being considered when coders are assigning ratings of global competence via these items. For instance, "overall competence" might be misconstrued as the therapist's competence exhibited throughout the session. In this case, aspects of technical competence (e.g., conducting psychoeducation in a developmentally appropriate manner) may be captured as a part of a therapist's overall skillfulness and responsiveness.

The only instrument that has been designed specifically to examine global competence in youth treatment is the Global Therapist Competence Scale for Youth Psychosocial Treatment (G-COMP; Brown et al., 2018). Brown et al. (2018) operationalized global competence with five domains (i.e., level-one items) that were all generated from previous global competence research and theory: (1) alliance building, (2) positive expectancies, (3) focusing treatment, (4) instigating change, and (5) responsiveness. The G-COMP provides examples of each of the five elements of

global competence (i.e., level-two items) and clearly differentiates global competence as separate from competencies specific to certain treatment programs. Thus, the specificity of this instrument's operational definition of global competence supports its use for assessing the construct in youth treatment. (*Note*: Brown et al. (2018) concluded that level-one items represented the most "parsimonious" approach to assessing global competence with the G-COMP. Hence, only level-one G-COMP items were considered in the current study and are hereafter referred to as "G-COMP items" or "G-COMP scores".)

In sum, the lack of a consistent definition of competence has created problems for measuring technical and global competence in youth treatment. It seems important for researchers studying the distinction between technical and global competence to utilize instruments that operationalize these constructs distinctly. The definitions that appear to best map onto the goal of informing improved training and evaluations of therapists are those that can provide more specific targets to be addressed in training and evaluation. Therefore, the following section focuses primarily on the psychometric properties of the CBAY-C and G-COMP, since these instruments are the two that most clearly defined technical and global competence, respectively.

Evaluation of Score Reliability and Validity of Competence Instruments

With definitions of technical and global competence provided, this section focuses on the score reliability and validity of these dimensions. Key psychometric properties to consider when evaluating competence instruments' scores in the youth treatment field include inter-rater reliability and construct validity. These psychometric properties are relevant to the aims of this study, as they allow research to determine if technical and global competence can be distinguished from one another in a reliable and valid way.

Inter-Rater Reliability

Inter-rater reliability refers to the degree to which multiple observers agree on the scores they generate when coding participants' performance on a given instrument (Kazdin, 2016). Inter-rater reliability is typically measured via intraclass correlation coefficients (ICC[2,2]s), which represent both the magnitude of correlations and the "agreement between measurements" (Koo & Li, 2016, p. 156). Because many competence instruments are observer-rated (including all the instruments previously reviewed), it is important to consider inter-rater reliability in the context of evaluating instruments of technical and global competence. Inter-rater reliability indicates the degree to which the data are "correct representations of the variables measured" (McHugh, 2012, p. 276) and provides information about how well coders agreed on scores of items that measure competence. If coders have low agreement on certain competence items, then it might suggest that those items are not able to capture the construct in a reliable way. Cicchetti (1994) proposed guidelines for assessing the level of clinical significance demonstrated by ICCs: ICCs below .4 are considered "poor", ICCs between .40 and .59 are considered "fair", ICCs between .60 and .74, are considered "good", and ICCs between .75 and 1.00 are considered "excellent".

Inter-rater reliability has been assessed for competence scores in the context of youth CBT sessions. For instance, Bjaastad et al. (2016) found inter-rater reliability that ranged from ICC(2,2) = .44 to .69, with a "fair" (Cicchetti, 1994) mean ICC(2,2) of .54 (*SD* = .12) for four technical competence items. Bjaastad et al. (2016) also found a "fair" ICC(2,2) of .49 for one global competence item in the context of CBT sessions for youth anxiety. Moreover, Hogue et al. (2008) found that ICC(2,2)s for technical competence items ranged from .01 to .63, with a "poor" (Cicchetti, 1994) mean ICC(2,2) of .35 (*SD* = .24). Hogue et al. (2008) also found that

ICC(2,2)s for global competence items ranged from .49 to .56, with a "fair" (Cicchetti, 1993) mean ICC(2,2) of .51 (*SD* = .04) in a CBT condition. Furthermore, inter-rater reliability was assessed for items on the CBAY-C, which operationalized technical competence most specifically and narrowly out of the instruments previously reviewed. Specifically, the CBAY-C model items have demonstrated ICC(2,2)s ranging from .37 to .80 (*M*=.65; *SD* = .12; McLeod et al., 2018). Similarly, inter-rater reliability has been examined for item scores on the G-COMP: Alliance Building: ICC(2,2) = .67, Positive Expectancies: ICC(2,2) = .70, Focusing Treatment: ICC(2,2) = .74, Instigating Change: ICC(2,2) = .79, Responsiveness: ICC(2,2) = .61 (*M* = .70, *SD* = .07; Brown et al., 2018). These findings suggest that CBAY-C model items and G-COMP items can be reliably assessed using two coders. While demonstrating evidence of inter-rater reliability is important, it is still necessary to examine other psychometric properties before drawing conclusions from competence instruments' scores.

Construct Validity

Another primary psychometric property to consider when assessing competence instruments' scores is construct validity, which is the extent to which scores on an instrument have been demonstrated to evaluate the construct of interest. Establishing construct validity necessitates multiple studies that produce results that are consistent and are what would be expected of the construct (Kazdin, 2016). Construct validity is established by "relating test scores to scores on measures of other theory-relevant variables" (Foster & Cone, 1995, pgs. 252-253). The construct validity of an instrument's scores cannot be reduced to a single correlation between scores on two instruments. Instead, construct validity involves the compilation of evidence from multiple sources, which can include other, related types of validity (Kazdin, 2016).

Specifically, convergent and discriminant validity should be evaluated when establishing the construct validity of an instrument's scores. Convergent validity refers to the degree to which scores on theoretically related instruments are correlated with one another. If two instruments are purported to measure the same construct, then it is expected that there will be a positive correlation between scores on those instruments (Kazdin, 2016). Thus, scores on technical and global competence instruments should be positively correlated with scores on instruments of related constructs (e.g., adherence and alliance; Brown et al., 2018; Hogue et al., 2008; McLeod et al., 2018). Conversely, discriminant validity is the degree to which scores on instruments of distinct constructs are not strongly correlated with one another (Kazdin, 2016). It is important to establish that scores on instruments of technical competence are distinct from scores on instruments of global competence if the two competence dimensions can be distinguished from one another. Hence, the current study evaluated convergent and discriminant validity of scores on technical and global competence instruments by assessing patterns of correlations between scores on competence instruments and scores on instruments of related constructs.

Because competence research is so sparse in the youth treatment field, there are few competence instruments to utilize in analyses of convergent and discriminant validity. Therefore, previous studies have utilized instruments of different, yet related constructs (e.g., adherence and alliance) to help establish the construct validity of competence scores. Previous studies (e.g., Brown et al., 2018; Hogue et al., 2008; McLeod et al., 2018) have conducted correlations between scores on their competence instruments and scores on instruments of related constructs as well as correlations between item and subscale scores within the same competence instrument. The construct validity of an instrument's scores is supported when the greatest magnitudes of correlations are found between those scores and scores on instruments assessing constructs that

are theoretically the most closely related. The following sections summarize the patterns of correlations between scores on instruments of technical and global competence and scores on theory-relevant instruments found in previous research. Notably, magnitudes of correlations were interpreted by using guidelines set forth by Rosenthal and Rosnow (1984): *r*'s = .10 to .23 were considered "small", *r*'s = .24 to .36 were considered "medium", and *r*'s = .36 or greater were considered "large". Multicollinearity was judged based on Kline (1979)'s guidelines, such that correlations above *r* =.7 indicated redundancy or conceptual overlap of items.

Competence

Scores on G-COMP items that assessed the five domains of global competence were found to be correlated with one another (*r'*s = .37 - .75, *M* = .51, *SD* = .11; Brown et al., 2018). Similarly, McLeod et al. (2018) found that scores on the CBAY-C global item that measured overall therapist skillfulness were correlated with scores on the standard items (*r'*s = .71 - .88; *M* =.81, *SD* = .09), model items (*r'*s = .59 - .91, *M* = .77, *SD* = .12) and delivery items (*r'*s = .69 - .86, *M* = .78, *SD* = .07). McLeod et al. (2018) also found that scores on the CBAY-C global item that measured overall therapist responsiveness were correlated with scores on the standard items (*r'*s = .73 - .89; *M* =.82, *SD* = .08), model items (*r'*s = .68 - .91, *M* = .80, *SD* = .08) and (delivery items: *r'*s = .72 - .87, *M* = .80, *SD* = .07). Moreover, Brown et al. (2018) found that G-COMP scores were associated with total scores on the CBAY-C: Alliance Building (*r* = .26), Positive Expectancies (*r* = .24), Focusing Treatment (*r* = .53), Instigating Change (*r* = .52), and Responsiveness (*r* = .26) (*M* = .36, *SD* = .15). Other studies have also examined the association between scores on technical and global competence items. For instance, Hogue et al. (2008) compared TBRS-C scores on the overall competence item to the goal average competence score (i.e., average score of individual technical competence items) in the CBT condition (*r* = .68) and

in the MDFT condition (r = .79). Additionally, Bjaastad et al. (2016) found a strong positive association between the scores on the CAS-CBT global rating of competence and the sum score of all four competence items (r = .88).

Overall, correlations amongst competence scores found in previous research can be classified as "large" (Rosenthal and Rosnow, 1984). Some correlations between scores of technical and global competence (e.g., Bjaastad et al., 2016; Hogue et al., 2008) can be classified as redundant (Kline, 1979). The only study that did not find that technical and global competence scores correlated at a level of redundancy was Brown et al. (2018); the mean inter-item correlation for G-COMP and CBAY-C scores (r = .36) was smaller than most other mean inter-item correlations found between technical and global competence item scores (e.g., Hogue et al., 2008; McLeod et al., 2018). These mixed patterns of correlations do not provide a clear answer as to whether scores on technical and global competence items can distinguish between the two dimensions of competence.

Adherence

Scores on competence instruments have also been compared to scores on instruments of adherence, which describes how closely the implementation of treatment matches the intended treatment plan (Barber et al., 2007; Southam-Gerow et al., 2016). For instance, correlations between CBAY-C scores and scores on corresponding items of the CBT Adherence Scale for Youth Anxiety (CBAY-A; Southam-Gerow et al., 2016) ranged from –.06 to .72 (M = .43, SD = .20; McLeod et al., 2018). The mean inter-item correlation would still be classified as "large" (Rosenthal & Rosnow, 1984), but not at a level that would suggest conceptual overlap between adherence and technical competence scores (Kline, 1979).

Similarly, global competence scores have also been found to be associated with adherence scores. For example, G-COMP scores were correlated with CBAY-A total scores: Alliance Building (r = .27), Positive Expectancies (r = .25), Focusing Treatment (r = .55), Instigating Change (r = .63), and Responsiveness (r = .26) (M = .39, SD = .18; Brown et al., 2018). The mean inter-item correlation between G-COMP scores and CBAY-A total scores was "large", but these results still suggested that global competence scores were distinct from adherence scores (Kline, 1979).

Hogue et al. (2008) also examined the discriminant validity of TBRS-C competence scores by comparing them to TBRS-C adherence scores. Goal average adherence scores (i.e., average score of all individual adherence items) were correlated with goal average competence scores in the CBT condition (r = .42) and in the MDFT (r = .17) condition. Overall competence scores were also correlated with adherence scores in CBT (r = .50) and MDFT (r = .23) condition. Likewise, Bjaastad et al. (2016) found a strong positive correlation between the sum scores on the adherence and competence items of the CAS-CBT (r = .79). Gutermann et al. (2015) also found that competence scores on the TCS were correlated with scores on an adherence instrument (Therapeutic Adherence Scale; r = .65).

While most of these studies found correlations between competence and adherence scores that were "large" (Rosenthal & Rosnow, 1984), only Bjaastad et al. (2016)'s results suggested that competence and adherence scores were redundant (Kline, 1979). Additionally, the mean correlation between G-COMP scores and CBAY-C total scores (M = .36; Brown et al., 2018) was smaller in magnitude than both the mean correlation between CBAY-C and CBAY-A scores (M = .43; McLeod et al., 2018) as well as the mean correlation between G-COMP and CBAY-A scores (M = .39; Brown et al., 2018). These mixed results bring the construct validity of CBAY-

C and G-COMP scores into question. If technical and global are both competence dimensions, then scores representative of these constructs should be more closely associated with one another than they are with scores of a separate construct (i.e., adherence). Thus, further examination of convergent and discriminant validity of scores on these instruments is needed to support their construct validity.

Alliance

Alliance is another construct theoretically related to competence that refers to the quality of the therapeutic relationship between the therapist and client (McLeod & Weisz, 2005). McLeod et al. (2018) compared CBAY-C scores to scores on the Therapy Process Observational Coding System-Alliance Scale (TPOCS-A; McLeod & Weisz, 2005). Correlations between CBAY-C and TPOCS-A scores ranged from .00 to .71 (M = .21, SD = .14; McLeod et al., 2018). Additionally, Brown et al. (2018) found the following correlations between scores on the last observations of G-COMP items and scores on an end-of-treatment youth-reported instrument of alliance called the Therapeutic Alliance Scale for Children (TASC; Shirk & Saiz, 1992): Alliance Building (r = .27), Positive Expectancies (r = .21), Focusing Treatment (r = .23), Instigating Change (r = .17), and Responsiveness (r = .22) (M = .22, SD = .04). Additionally, correlations between G-COMP scores and TPOCS-A scores were also conducted: Alliance Building (r = .44), Positive Expectancies (r = .26), Focusing Treatment (r = .27), Instigating Change (r = .43), and Responsiveness (r =.18) (M= .32, SD = .11; Brown et al., 2018). Finally, Hogue et al. (2008) found that scores on an instrument of child-therapist alliance, the Vanderbilt Therapeutic Alliance Scale-Revised (VTAS-R; Hogue et al., 2006) were not significantly related to goal average competence scores in the CBT subscale but were correlated with goal average competence scores in the MDFT condition (r = .4). Moreover, VTAS-R scores were correlated

with TBRS-C "overall" competence scores in the CBT condition (r = .31) and in the MDFT condition (r = .36).

Most mean correlations found in previous studies between competence and alliance scores were "small" (Rosenthal & Rosnow, 1984). Overall, correlations between competence and adherence scores were larger in magnitude than correlations between competence and alliance scores. For instance, the mean correlation between CBAY-C and CBAY-A scores (r = .43; McLeod et al., 2018) was larger than the mean correlation between CBAY-C and TPOCS-A scores (r = .21; McLeod et al., 2018). Similarly, the mean correlation between G-COMP and the CBAY-A scores (r = .39; Brown et al., 2018) was larger than the correlations between G-COMP and TPOCS-A scores (r = .32; Brown et al., 2018) and between the last observations of G-COMP and TASC scores (r = .22; Brown et al., 2018). The non-redundant (Kline, 1979), "small" (Rosenthal & Rosnow, 1984) magnitudes of correlations between G-COMP and TASC scores, between G-COMP and TPOCS-A scores, as well as between CBAY-C and TPOCS-A scores suggest that the G-COMP and CBAY-C items are likely measuring a construct that is not alliance. Additionally, the construct validity of the CBAY-C and G-COMP scores are supported by the magnitudes of correlations between scores on these instruments and alliance scores (TPOCS-A and TASC) being smaller than the magnitudes of correlations between these scores and adherence scores (CBAY-A).

Summary

McLeod et al. (2018) and Brown et al. (2018) examined several psychometric properties of scores on instruments of technical (CBAY-C) and global (G-COMP) competence designed for use in youth treatment. Their analyses of inter-rater reliability suggested that items on these

instruments can be reliably assessed using two coders. However, the construct validity of CBAY-C and G-COMP competence scores is less clear.

In the extant literature, the largest correlations were generally found between scores of technical and global competence. Some associations were large enough to suggest that these two constructs were not distinct from one another (e.g., Bjaastad et al., 2016; Hogue et al., 2008), while other findings suggested these two dimensions of competence can be distinguished from one another (i.e., Brown et al., 2018). Notably, Brown et al. (2018) posited that the use of independent coders could have contributed to the conceptual distinction between global and technical competence in their study. Brown et al. (2018) also conducted correlations between scores on the G-COMP items and the CBAY-C *total* score. Since the CBAY-C total score included items that might not completely isolate technical competence (e.g., standard and global items), Brown et al. (2018)'s findings may not be representative of the association between technical and global competence as defined by the current study. Inconsistency in both definitions and methods of measuring these constructs is a likely reason for the unclear findings regarding whether technical and global competence are distinct constructs (McLeod et al., 2018).

Another point of uncertainty was that correlations between CBAY-C and G-COMP scores were smaller than both correlations between CBAY-C and CBAY-A scores as well as correlations between G-COMP and CBAY-A scores. Since technical and global competence are both dimensions of competence, it would have been expected that technical (CBAY-C) and global (G-COMP) competence scores would be more strongly associated with each other than with adherence (CBAY-A) scores (which was a pattern found by Hogue et al., 2008). The pattern of correlations between CBAY-C, CBAY-A, and G-COMP scores brings the construct validity of CBAY-C and G-COMP scores into question. Therefore, further investigation is needed to

examine the convergent and discriminant validity of CBAY-C and G-COMP scores to inform conclusions about the construct validity of scores on these instruments. The construct validity of CBAY-C and G-COMP scores is important to establish before meaningful conclusions can be drawn from these instruments' scores about the distinction between the two dimensions of competence.

A clearer pattern emerged with regards to competence and alliance scores. Since most correlations between competence (CBAY-C and G-COMP) and alliance (TPOCS-A and TASC) scores have been smaller than those between competence (CBAY-C and G-COMP) and adherence (CBAY-A) scores, CBAY-C and G-COMP scores can be understood to be assessing constructs distinct from those assessed by the TPOCS-A and TASC. However, an important caveat is that discriminant validity of scores does not automatically solidify the construct validity of these scores. For instance, it might be reasonable to conclude that scores on instruments of competence and alliance are assessing distinct constructs. Yet, it cannot be definitively concluded that scores on these instruments are assessing the constructs they purport to assess. Thus, the small correlations between CBAY-C and TPOCS-A scores as well as between G-COMP and TPOCS-A/TASC scores does not automatically solidify the construct validity of CBAY-C and G-COMP scores.

Thus, the current literature does not provide enough evidence to suggest that technical and global competence, as they are defined in the current study, can be measured distinctly. It is important to remedy this confusion because understanding these two constructs individually could potentially improve the ways in which therapists are trained and evaluated.

Present Study

If technical and global competence are to be understood and utilized as distinct dimensions of competence, there needs to be more empirical evidence that suggests that they are distinct dimensions that can be measured separately. Yet, clarifying the definitions of technical and global competence appears to be a prerequisite for distinguishing between the two dimensions of competence. In the present study, *technical competence* was defined as a therapist's skillfulness and responsiveness when implementing a specific technique (i.e., practice element) found in a particular treatment program (Barber et al., 2007; Hogue et al., 2008; McLeod et al., 2018). Conversely, *global competence* was defined as a therapists' general "clinical acumen" that permeates their interventions (Barber et al., 2007; Sharpless & Barber, 2009) and cuts across various treatment modalities, interventions and programs (Brown et al., 2018).

The primary aim of the current study was to determine if technical and global competence can be measured as distinct dimensions of competence in CBT programs for youth anxiety. To meet this aim, the current study utilized data from a randomized effectiveness trial (Child STEPs; Weisz et al., 2012), which is described below in greater detail. The instruments used to assess competence of Child STEPs therapists (i.e., CBAY-C and G-COMP) allowed the current study to address its primary aim, because these instruments mapped onto the provided definitions of technical and global competence. The CBAY-C model items were used as a measure of technical competence, as they assess the competence associated with the implementation of discrete practice elements. Additionally, based on previous research regarding common elements that should be found in all treatments (e.g., Castonguay & Beutler; Frank, 1971), the items of the G-COMP (i.e., Alliance Building, Positive Expectancies, Focusing

Treatment, Instigating Change, and Responsiveness) were used to assess global competence. The G-COMP items fit the provided definition of global competence, as they describe certain clinical skills that are not bound to any one treatment program but rather cut across various therapeutic interactions.

The sample and design of the Child STEPs study was suitable to address the primary aim of the current study for several reasons. First, Child STEPs child participants were all aged 7 to 13, which was an appropriate sample for studying competence in youth treatment. Second, 31.6% of Child STEPs child participants had a presenting problem of anxiety that was treated with CBT, which allowed the current study to examine competence in the context of CBT for youth anxiety. Therefore, the current study only utilized this subsample of Child STEPs.

The Child STEPs subsample of children being treated with CBT for anxiety was also a good match for the current study's instruments of interest. For instance, the CBAY-C was designed specifically to assess technical competence in the context of CBT for youth anxiety (McLeod et al., 2018); hence, it was appropriate to use with the subsample of Child STEPs participants who had a primary presenting problem of anxiety. Moreover, the G-COMP was designed to measure global competence across various modalities of youth treatment. The psychometric properties of the G-COMP were also first assessed by utilizing a sample of youth receiving CBT for anxiety (Brown et al., 2018). Thus, the Child STEPs sample was an appropriate sample with which to utilize the G-COMP. Because the Child STEPs sample was a good fit for the current study's population and instruments of interest, the current study was well positioned to examine the distinction between technical and global competence in the context of CBT for youth anxiety. Also, the Child STEPS sample was a different youth sample from the

one used in Brown et al. (2018)'s study; thus, the current study added a unique datapoint towards the goal of discriminating between technical and global competence in youth treatment.

The design of Child STEPs was also appropriate to address the primary aims of the current study. Child STEPS was comprised of three conditions: modular-manualized treatment (MMT), standard manualized treatment (SMT), and usual care. Therapists in both the MMT and SMT conditions implemented CBT for youth anxiety. However, therapists in the SMT condition followed a pre-determined course of treatment; whereas, therapists in the MMT condition were provided flowcharts with suggested courses of treatment but were free to change course depending on the needs of their individual clients. Additionally, therapists in the MMT condition were able to draw upon modules meant for other problem areas (e.g., depression or conduct) to treat their clients who originally presented with a primary problem of anxiety when appropriate. Hence, treatment in the SMT condition was more structured, while treatment in the MMT condition was able to be implemented in a more flexible manner.

Therapists in the SMT and MMT groups implemented CBT for youth anxiety, which provided the opportunity to assess both global competence (via the G-COMP) and technical competence for CBT for youth anxiety (via the CBAY-C). However, because the nature of the MMT condition was more flexible than the structure of the SMT condition, therapists in the MMT condition may have had more opportunities to display global competencies related to flexibility in treatment (e.g., responsiveness). Conversely, it may have been easier for SMT therapists to display technical competencies because of the more highly structured nature of the SMT condition. Moreover, there were certain practice elements present in the SMT condition that were not present in the MMT condition and vice versa. Because the implementation of

treatment was different between the MMT and SMT conditions, data from these two conditions were analyzed separately in the current study.

Additionally, because therapists in the usual care condition did not consistently implement CBT, only scores from the MMT and SMT conditions were used in analyses. Since technical competence is defined as skillfulness and responsiveness in the implementation of practice elements of a particular program, and "usual care" is not a specific treatment program, it would not have been appropriate to measure technical competence in this condition. Similarly, the usual care condition would likely not have provided sufficient demonstrations of technical competence for CBT for youth anxiety, which is necessary when assessing technical competence with the CBAY-C.

Furthermore, how treatment processes of the Child STEPs sessions were assessed allowed the current study to evaluate the construct validity of the CBAY-C and G-COMP scores. First, all treatment sessions were double-coded, which allowed for the examination of inter-rater reliability of the CBAY-C and G-COMP scores. Additionally, the instruments used to assess the treatment integrity (i.e., CBAY-C, G-COMP, CBAY-A) and related constructs (i.e., TPOCS-A) of Child STEPs sessions were appropriate for analyses of convergent and discriminant validity of CBAY-C and G-COMP scores. The current study assessed the convergent and discriminant validity of CBAY-C and G-COMP scores in a manner similar to previous research (i.e., Brown et al., 2018; Hogue et al., 2008; McLeod et al., 2018), which has assessed the construct validity of competence scores by comparing them to scores of instruments that assess similar constructs (i.e., adherence and alliance). Specifically, the current study used patterns of correlations amongst scores of these constructs to draw preliminary conclusions about the construct validity of CBAY-C and G-COMP scores.

The following hypotheses were tested in the current study:

Inter-rater reliability:

1. *Hypothesis 1*. It was hypothesized that scores on applicable CBAY-C model items would demonstrate fair to excellent inter-rater reliability (ICC(2,2) ≥ .40; Cicchetti, 1994) in both conditions, with a mean ICC(2,2) in the "good" range (ICC(2,2) ≥ .60; Cicchetti, 1994), as found in previous research (McLeod et al., 2018).
2. *Hypothesis 2*. It was hypothesized that scores on G-COMP items (i.e., the five dimensions of global competence) would demonstrate good to excellent inter-rater reliability (ICC(2,2) ≥ .60; Cicchetti, 1994) in both conditions, with a mean ICC(2,2) in the "good" range (ICC(2,2) ≥ .60; Cicchetti, 1994), as found in previous research (Brown et al., 2018).

Construct validity

3. *Hypothesis 3*. It was hypothesized that correlations between the CBAY-C model and G-COMP scores would be positive and "large" (*r*'s ≥ .36; Rosenthal & Rosnow, 1984), but not at a level that would suggest redundancy in both conditions (*r*'s ≤ .7; Kline, 1979), which would be similar to findings of previous research (Brown et al., 2018).
4. *Hypothesis 4*. It was hypothesized that CBAY-C model scores as well as G-COMP scores would be positively correlated with scores on an instrument of adherence (i.e., CBAY-A) at a level classified as "large" (*r*'s ≥ .36; Rosenthal & Rosnow, 1984) in both conditions. The magnitudes of these correlations were expected to be smaller than the magnitudes of correlations between CBAY-C model scores and G-COMP scores in both conditions, similar to patterns found in previous research (e.g., Hogue et al., 2008).

5. *Hypothesis 5*. It was hypothesized that CBAY-C model scores as well as the G-COMP scores would be positively correlated with TPOCS-A scores at a "small" to "medium" level ($.10 \leq r \leq .24$; Rosenthal & Rosnow, 1984) in both conditions, similar to findings demonstrated in previous research (e.g., McLeod et al., 2018; Brown et al., 2018). The magnitudes of these correlations were expected to be smaller than the magnitudes of correlations between scores on competence instruments (CBAY-C and G-COMP) and scores on an adherence instrument (CBAY-A) in both conditions, as demonstrated in previous research (e.g., McLeod et al., 2018; Brown et al., 2018).

Method

Data Source

The current study utilized data from a randomized effectiveness trial (Child STEPs; Weisz et al., 2012). Child STEPs was conducted in 10 outpatient clinics in Massachusetts and Hawaii. The primary aim of Child STEPs was to compare the effectiveness of two treatment program designs in treating youth anxiety, depression, or conduct problems. Child STEPs therapists were randomly assigned to one of three conditions: (1) modular manualized treatment, (2) standard manualized treatment, or (3) usual care treatment.

Participants

Youth participants. There were several inclusion criteria for youth participants in Child STEPs. First, participants had to meet criteria for one of the following Diagnostic and Statistical Manual of Mental Disorders, 4th ed. (DSM-IV-TR; American Psychiatric Association, 2000) disorders: anxiety, depression, or conduct problems, as determined by the Children's Interview for Psychiatric Symptoms (Weller et al., 2000) or had to demonstrate clinically elevated problem levels ($T > 65$) in one or more of the three problem areas, as measured by the Child Behavior

Checklist and the Youth Self Report (Achenbach & Rescorla, 2001). Additionally, the participants' families needed to seek treatment (i.e., recruitment was not conducted via advertisements). Exclusion criteria for Child STEPs included: (1) intellectual disability, pervasive developmental disorder, psychotic symptoms, or bipolar disorder; or (2) a primary problem of inattention or hyperactivity. Ultimately, 174 youths ages 7 to 13 (M = 10.59 years, SD = 1.76) participated in Child STEPs. Yet, the current study included only youth participants whose primary presenting problem was anxiety (31.6% of the Child STEPs sample, n = 38) and assigned to one of the two manualized treatment conditions. Additionally, youth participants were only included in the current study if they had at least two audible treatment sessions and received treatment from only one therapist. There were 38 child participants in the current study who ranged in age from 8 to 13 years old (M = 9.84, SD = 1.65) and who identified as the following: 60.5% non-Hispanic White, 2.6% Asian, 5.3% Black of African American, 2.6% Latinx, 26.3% mixed race, and 2.6% other. There were 20 male (52.6%) and 18 (47.4%) female child participants in the current study (see Results and Tables 2 and 3 for further detail).

Therapist participants. Child STEPs included 84 therapists from 10 outpatient clinics in Massachusetts and Hawaii. These therapist participants (M age = 40.35 years, SD age = 9.67) delivered either modular manualized treatment (MMT), standard manualized treatment (SMT), or usual care. Treatment sessions were conducted in either a school-based or outpatient community service setting. The current study only included therapists that delivered individual treatment sessions to youth with a primary presenting problem of anxiety in either the MMT or SMT group. There were 26 therapist participants in the current study who ranged in age from 27 to 59 years old (M = 40.34, SD = 9.67). The following list details the gender and racial identities of the therapist participants in the current study: 80.8% were female and 19.2% were male.

Moreover, 53.8% were Caucasian, 23.1% were Asian American, 7.7% were African American, and 7.7% identified as "other"; ethnicity was not reported for 7.7% of the sample (see Results and Tables 2 and 3 for further detail).

Treatment Conditions

The Child STEPs therapists were randomly assigned to either the MMT, SMT or usual care group based on a cluster randomization design (Campbell, Elbourne, & Altman, 2004). In this system, a blocked randomization stratified by therapist educational level (master's vs doctoral degree) was utilized, and the allocation ratio for each block was 1:1:1. Youth participants and their families were informed that they would be randomly assigned into a treatment group; however, they were all naïve to the treatment condition to which they were assigned.

Modular manualized treatment. Therapists in the MMT condition of Child STEPs utilized the *Modular Approach to Therapy for Children with Anxiety, Depression, and Conduct Problems* (*MATCH*; Chorpita & Weisz, 2005) protocol. MATCH is comprised of treatment modules that address anxiety, depression, and conduct problems. The modules in MATCH correspond to treatment procedures delivered in the Coping Cat, PASCET, and Defiant Children (see below for detail) protocols. There were flowcharts for each problem area in MATCH that detailed a default sequence of modules. Therapists in the MMT condition chose the flowchart associated with the primary problem area identified as the highest treatment priority by the youth and their caregiver (assessed via the Top Problems Assessment; Weisz et al., 2011). If a crisis, stressor or comorbid condition arose during the course of treatment, the therapists in the MMT condition could deviate from the flowchart by incorporating modules from another flowchart or

by switching to another flowchart entirely. There were 16 children in the MMT condition with a primary presenting problem of anxiety.

Standard manualized treatment. The SMT condition of Child STEPs was comprised of three treatment protocols with manualized instructions and a prescribed order of treatment sessions. The following protocols were utilized: (1) Coping Cat, which is an individual cognitive behavioral therapy (ICBT) protocol for anxiety (Kendall, 1994; Kendall & Hedtke, 2006a, 2006b), (2) Primary and Secondary Control Enhancement Training (PASCET), which is an ICBT protocol for depression (Weisz, Weersing, Valeri, & McCarty, 1999), and (3) Defiant Children, which is a behavioral parent training protocol for conduct problems (Barkley, 1997). Therapists administered the Top Problems Assessment (Weisz et al., 2011) to decide whether to begin treatment with a focus on anxiety, depression, or conduct problems. Therapists first delivered Coping Cat if the primary presenting problem was anxiety. Therapists first delivered PASCET if the primary presenting problem was depression. Lastly, therapists first delivered Defiant Children if the primary presenting problem was conduct.

As previously mentioned, the current study only included youth with a primary presenting problem area of anxiety. Thus, the current study only included therapists who delivered Coping Cat first. Coping Cat is comprised of 16-20 sessions that are designed to address anxiety symptomology through skill-building (e.g., cognitive restructuring, relaxation, problem solving), graduated exposure to feared stimuli or situations, and continued practice of skills both in (e.g., role plays) and out (i.e., homework assignments) of the treatment sessions. There were 21 children in the SMT condition with a primary presenting problem of anxiety.

Therapist training and consultation. Child STEPS therapists in both the MMT and SMT conditions participated in six days of training together; two days of training were

designated for each of the three problem areas. Therapists in the MMT and SMT conditions both received weekly consultation on cases from supervisors. The Child STEPs study utilized a feedback system that allowed consultants to track the delivery of treatment practices in order to increase adherence to the study's protocols. In each feedback sessions, supervisors used a checklist of treatment practices to ask therapists about what content they delivered (e.g., relaxation), and the techniques they used to deliver that content (e.g., role play). Supervisors gathered information from therapists in the MMT and SMT conditions and then provided guidance and support to those therapists. The supervisors also discussed measurement feedback and client progress with the therapist participants.

Instruments

The Cognitive-Behavioral Treatment for Anxiety in Youth Competence Scale

(CBAY-C; McLeod et al., 2018) is an observational instrument designed to capture technical competence (i.e., skillfulness and responsiveness specific to a particular treatment program) in the delivery of core practice elements found in ICBT for youth anxiety. The original CBAY-C was comprised of 25 total items: five Standard (interventions that are commonly used in CBT programs but are not unique to ICBT for youth anxiety, such as reviewing homework), 12 Model (core theory-driven interventions specific to ICBT for youth anxiety and that should be the focus of one or more sessions, such as relaxation), six Delivery (the way in which model interventions are delivered, such as rehearsal), and two Global (overall skillfulness and responsiveness). Scores on the original CBAY-C have shown evidence of inter-rater reliability at the item level (ICC[2,2]s ranged from .37 to .80; M = .67, SD = .11), as well as representativeness, convergent and discriminant validity (see McLeod et al., 2018).

The CBAY-C was adapted for use in the current study. The adapted version was comprised of 21 of the initial items (two Standard items and the two Global items were removed) as well as 12 new items that correspond with the new CBAY-A items, for a total of 33 items. In the current study, coders produced competence scores on the CBAY-C by considering both skillfulness (i.e., quality of the delivery) and responsiveness (i.e., the timing and appropriateness of delivery). More specifically, coders were asked to consider four dimensions of skillfulness and responsiveness when producing competence scores for each item: "(a) expertise, commitment, motivation; (b) clarity of communication; (c) appropriate timing of delivery; and (d) read and respond to where the client appears to be regarding level of therapeutic engagement/ understanding" (McLeod et al., 2018, p. 51). Competence scores were produced from a 7-point Likert-style scale with the following anchors: 1 (very poor), 3 (acceptable), 5 (good), and 7 (excellent). Competence scores of zero were given when the Model, Delivery or Standard item was not present in the session (i.e., when adherence for the corresponding item was scored a 1).

The Global Therapist Competence Scale for Youth Psychosocial Treatment (G-COMP; Brown et al., 2018) is an observational instrument designed to measure global competence (i.e., skillfulness and responsiveness that should be present in all therapists, regardless of what particular treatment program they are delivering) in youth treatment. The G-COMP is comprised of five items: Alliance Building, Positive Expectancies, Focusing Treatment, Instigating Change, and Responsiveness. Scores on the G-COMP have previously demonstrated inter-rater reliability (ICC[2,2]s > .60; M = .70, SD = .07) and construct validity (Brown et al., 2018). Coders produced competence ratings of global competence on a 7-point Likert-style scale from 1 (very poor) to 7 (excellent) (Brown et al., 2018).

The Cognitive Behavioral Therapy Adherence Scale for Youth Anxiety (CBAY-A; Southam-Gerow et al., 2016) is an observational instrument designed to capture adherence to common practice elements found in ICBT for youth anxiety. The CBAY-A was also adapted for use in the current study. The adapted version was comprised of 35 items: six Standard (i.e., prescribed interventions common to many CBT programs, such as homework assignment), 22 Model (i.e., interventions specific to ICBT for youth anxiety, such as exposure), and seven Delivery (i.e., the way in which model interventions are delivered, such as rehearsal). Scores on the original, 23-item version of the CBAY-A have previously demonstrated evidence of item- and subscale-level score reliability (ICC[2,2]s ranged from .48 to .80; M = .77, SD = .15) and construct validity for use with an ICBT protocol (i.e., Coping Cat; Kendall & Hedtke, 2006a, 2006b; Southam-Gerow et al., 2016). Adherence was scored on a 7-point Likert-style scale with the following anchors: 1 = not at all, 3 = somewhat, 5 = considerably, and 7 = extensively. When coding adherence with the CBAY-A, coders considered both the frequency (i.e., how often an item was observed) and thoroughness (i.e., the amount of time and effort spent on an item across the session) with which a therapist implemented an intervention.

The Therapy Process Observational Coding System-Alliance Scale (TPOCS-A; McLeod & Weisz, 2005) is an observational instrument of the youth-therapist alliance. The TPOCS-A is comprised of nine items that are divided into two categories: bond (i.e., affective aspects of the youth-therapist relationship) and task (i.e., client participation in the activities of treatment). The bond and task categories represent two commonly emphasized dimensions of alliance (Shirk & Russell, 1998). Scores on the TPOCS-A have previously demonstrated evidence of inter-rater reliability (ICC[2,2]s ranged from .40 to .75; M = .59, SD = .10),

convergent validity with a self-report alliance instrument (TASC; Shirk & Saiz, 1992), and predictive validity with child outcomes (Liber et al., 2010; McLeod & Weisz, 2005).

Study Procedures

Coding procedures. One coding team was created for the CBAY-C and G-COMP, and another was created for the CBAY-A and TPOCS-A. Principal investigators trained coding teams separately over the course of approximately three months until an adequate item-level of reliability was reached (ICC[2,2] ≥ .60). The CBAY-C and G-COMP coding team was comprised of two female clinical psychology doctoral students (50% Latinx, 50% White non-Latinx). The CBAY-A and TPOCS-A coding team was comprised of three female clinical psychology doctoral students (33% Asian-American; 66% White non-Latinx). One coder served on both teams. Coders were blind to the treatment condition (i.e., MMT, SMT or usual care) and coded sessions in an order randomly assigned by the principal investigators.

In the beginning stages of training, coders read and discussed the scoring manual. Principal investigators also reviewed coded sessions with the coders. Then, coders coded recordings of treatment sessions independently and participated in weekly meetings in which results of the practice coding were discussed. Lastly, coders began coding treatment sessions for the certification phase, during which they were required to reach an adequate level of reliability across 32 recordings (ICC[2,2] ≥ .60). Treatment sessions were recorded as audio and video files.

After the coders reached an adequate level of reliability and were considered "certified coders", they began independently coding randomly assigned sessions. Principal investigators and coders met regularly throughout the independent coding phase to prevent coder drift (Margolin et al., 1998), which was assessed via the continued examination inter-rater reliability

(i.e. reliability coefficients). If an item fell below the acceptable level of reliability (i.e., ICC[2,2] < .60), additional training was provided (e.g., reexamination of the coding manual, group discussions about the discrepancies, and/or group coding of problematic items). Notably, coders were blind to treatment condition. That is, coders were unaware if a particular session was an MMT, SMT or usual care session.

Sampling of treatment sessions. Except for the first and last sessions for each client, all available treatment sessions were selected from each client for coding and randomly assigned to coders. The final sample of sessions coded with the CBAY-C and G-COMP consisted of 796 coded sessions (244 in MMT, 359 in SMT, and 193 sessions in usual care). The current study only used recordings of therapists in the MMT and SMT conditions, for a total of 603 coded recordings.

Data Analysis Plan

The purpose of the current study was to determine if technical competence and global competence could be measured as distinct constructs in the context of CBT for youth anxiety via scores on the CBAY-C model items and G-COMP items. Inter-rater reliability and construct validity were examined for technical (CBAY-C model) as well as global competence (G-COMP) scores.

Data were compared on key demographic and clinical characteristics (e.g., race/ethnicity, age, sex, Child Behavior Checklist or CBCL scores) to examine any differences between the two conditions (i.e., MMT or SMT). Patterns of missing data were evaluated to determine if data were missing completely at random (MCAR), missing at random (MAR), or missing not a random (MNAR; Schafer & Graham, 2002). The mean, standard deviation, range, and

distribution (i.e., skewness and kurtosis) were examined for technical competence (CBAY-C model) item and subscale scores as well as for global competence (G-COMP) subscale scores.

Subscale Generation and Inter-Rater Reliability

The subscales used in analyses only included items that aligned with the definitions of each dimension of competence previously provided; hence, CBAY-C model items were used as measures of technical competence and G-COMP items were used as measures of global competence. All analyses were conducted separately for the MMT and SMT conditions. Besides the overall structures of the two treatment conditions being different, there were key differences in content across the two conditions. There were nine technical competence (CBAY-C model) items that were applicable to the MMT condition and 11 technical competence (CBAY-C model) items that were applicable to the SMT condition (see Table 1). To match the CBAY-C model items to each condition, CBT experts categorized items as either applicable to MMT (n = 9 items) or SMT (n = 11 items) based on content of the MATCH (MMT) and Coping Cat (SMT) manuals, respectively.

As the primary aim of the current study was to determine whether technical and global competence can be measured as distinct constructs in CBT programs for youth anxiety, it was appropriate to only include technical competence items that aligned with the specific treatment program for which they were applicable. Hence, MMT and SMT analyses were conducted separately. The CBAY-C MMT subscale included the following technical competence items: 1) Psychoeducation, 2) Emotion Education, 3) Fear Ladder, 4) Cognitive Anxiety, 5) Coping Plan, 6) Exposure: Prep, 7) Exposure, 8) Exposure: Debrief, and 9) Maintenance. Whereas, the CBAY-C SMT subscale included the following technical competence items: 1) Psychoeducation, 2) Emotion Education, 3) Fear Ladder, 4) Relaxation, 5) Cognitive Anxiety, 6) Problem Solving,

7) Self-Reward, 8) Coping Plan, 9) Exposure: Prep, 10) Exposure, and 11) Exposure: Debrief. Similarly, CBAY-A model subscales were created by matching the content of the CBAY-A model items to the content of the conditions' protocols (i.e., MATCH or Coping Cat). That is, the CBAY-A MMT subscale contained the same items as the CBAY-C MMT subscale, and the CBAY-A SMT subscale contained the same items as the CBAY-C SMT subscale.

Conversely, the global competence (G-COMP) items were equally applicable to both the MMT and SMT conditions. Thus, global competence (G-COMP) subscales were identical across conditions and were comprised of the G-COMP items, as was suggested by previous research (Brown et al., 2018).

Furthermore, inter-rater reliability was considered when generating subscales. The inter-rater reliability of technical (CBAY-C model) and global competence (G-COMP) items was evaluated by estimating intra-class correlation coefficients (ICCs; Shrout & Fleiss, 1979). The model ICC(2,2) based on a two-way random effects model was used, as it provides a reliability estimate of the average score of the coders and allows for generalizability of the findings to other samples (Shrout & Fleiss, 1979). Cicchetti (1994)'s guidelines were used to evaluate the value of the ICC(2,2)s: below .40 was considered "poor", between .40 and .59 was considered "fair", between .60 and .74 was considered "good", and .75 and above were considered "excellent". Two hypotheses related to inter-rater reliability were evaluated. Hypothesis 1 predicted that CBAY-C model item scores would demonstrate at least "fair" inter-rater reliability (ICC[2,2] ≥ .4) in both conditions. Similarly, hypothesis 2 predicted G -COMP item scores would demonstrate at least "fair" inter-rater reliability (ICC[2,2] ≥ .4) in both conditions. Competence items were considered for exclusion from subscales if the items demonstrated "poor" (Cicchetti, 1994) inter-rater reliability.

Construct Validity: Technical Competence

To evaluate the construct validity of two competence instruments (CBAY-C and G-COMP), correlations were first conducted to evaluate the overlap between technical (CBAY-C model) and global competence (G-COMP) scores in the MMT and SMT conditions. Per Hypothesis 3, in both conditions, there were expected to be "large" (Rosenthal & Rosnow, 1984) positive correlations between technical (CBAY-C model) and global competence (G-COMP) scores, but not at a level that would suggest redundancy (Kline, 1979). These results would support the discriminant validity of CBAY-C model and G-COMP scores, such that these scores would be measuring distinct constructs in two ICBT programs for youth anxiety.

Then, bivariate correlations were conducted to examine the associations between technical competence (CBAY-C model) and adherence (CBAY-A model) scores in the MMT and SMT conditions. Per Hypothesis 4, in both conditions, there were expected to be "large" (Rosenthal & Rosnow, 1984) positive correlations between technical competence (CBAY-C model) and adherence (CBAY-A model) scores. The magnitudes of these correlations were expected to be smaller than the magnitudes of the correlations between technical (CBAY-C model) and global competence (G-COMP) scores.

Next, bivariate correlations were conducted to examine the associations between technical competence (CBAY-C model) and alliance (TPOCS-A) scores in the MMT and SMT conditions. Per Hypothesis 5, in both conditions, there were expected to be "small" to "medium" (Rosenthal & Rosnow, 1984) positive correlations between technical competence (CBAY-C model) and alliance (TPOCS-A) scores. The magnitudes of these correlations were expected to be smaller than the magnitudes of the correlations between technical competence (CBAY-C model) and adherence (CBAY-A model) scores. This pattern of correlations would be similar to

those found in previous research (e.g., Hogue et al., 2008; McLeod et al., 2018) and support the construct validity of the technical competence (CBAY-C model) scores. Additionally, follow-up contrasts were examined using Fisher *r*-to-*z* transformation.

Construct Validity: Global Competence

Bivariate correlations were also conducted to examine the association between global competence (G-COMP) and adherence (CBAY-A model) scores in the MMT and SMT conditions. Per Hypothesis 4, in both conditions, there were expected to be "large" (Rosenthal & Rosnow, 1984) positive correlations between global competence (G-COMP) and adherence (CBAY-A model) scores. The magnitudes of these correlations were expected to be smaller than the magnitudes of the correlations between technical competence (CBAY-C model) and global competence (G-COMP) scores.

Then, bivariate correlations were conducted to examine the association between global competence (G-COMP) and alliance (TPOCS-A) scores in the MMT and SMT conditions. Per Hypothesis 5, in both conditions, there were expected to be "small" to "medium" (Rosenthal & Rosnow, 1984) positive correlations between global competence (G-COMP) and alliance (TPOCS-A) scores. The magnitudes of these correlations were expected to be smaller than the magnitudes of the correlations between global competence (G-COMP) and adherence (CBAY-A model) scores. This pattern of correlations would be similar to those found in previous research (e.g., McLeod et al., 2018; Brown et al., 2018) and would support the construct validity of the global competence (G-COMP) scores. Moreover, follow-up contrasts were examined using Fisher *r*-to-*z* transformation.

Lastly, post-hoc principal components analyses were conducted to examine technical and global competence scores in both study conditions. Principal components analyses were

considered the most appropriate data reduction techniques (as opposed to factor analyses), as the goal of the study was not to examine latent constructs of the current data. Instead, the primary aim of the current study was to determine whether technical and global competence could be measured as distinct domains of competence via the CBAY-C model and G-COMP items. Since principal component analyses reduce data to create index variables (or components) to explain the total variance, the use of principal component analyses was best suited for the primary aim of the current study. Moreover, the exploratory nature of principal component analyses was appropriate for the current study's aim, as there is currently no theoretical consensus regarding whether technical and global competence can be measured distinctly. Technical competence (i.e., CBAY-C MMT or CBAY-C SMT) scores and global competence scores (i.e., G-COMP: Alliance Building, G-COMP: Positive Expectancies, G-COMP: Focusing Treatment, G-COMP: Instigating Change, and G-COMP: Responsiveness) were entered into analyses. Because components were expected to be correlated, competence scores were analyzed using principal component analysis with Oblique (non-orthogonal) rotation (DeVellis, 2017) and components were extracted if eigenvalues were greater than one (Kaiser, 1960).

Results

Demographic Data

Comparisons between study groups and between those included and excluded from analyses were conducted to examine group differences in the current sample (i.e., the anxiety subsample of Child STEPs). There were six youth participants who were excluded from the current study's analyses because they had fewer than three recorded sessions. Youth participants who were excluded from analyses did not differ from the current sample in any of the key demographic or clinical variables (see Table 2).

Demographic and Child Behavior Checklist (CBCL) data were also compared between the participants in the modular manualized treatment (MMT) and standard manualized treatment (SMT) conditions. Youth participants in the MMT condition did not differ from those in the SMT condition on any of the key demographic or clinical variables (see Table 3). Demographic and clinical data were also compared between therapists in the two conditions. There was only was significant difference; the mean age of therapists was significantly lower in the MMT condition (M = 35.20, SD = 6.81) than the SMT condition (M = 43.56, SD = 9.96), $t(24) = -2.33$, $p = .03$.

Missing Data Analyses

Technical (CBAY-C) and global (G-COMP) competence instruments were used to code a subset of treatment sessions from the Child STEPs study. To examine whether the coded MMT and SMT sessions were representative of overall treatment, the proportion of sessions coded was compared between the two groups using an independent samples t-test. There was not a significant difference in percent of sessions coded between the MMT group (M = .74, SD = .19) and the SMT group (M = .74, SD = .13; $t(36) = -.06$, $p = .95$). A total of 4% of therapist-level demographic information (i.e., race/ethnicity information for two therapists, years' experience for two therapists, specialty area for one therapist) were missing, and was considered to be MCAR (Little's MCAR test Chi Square = 19.67, df = 15, $p = .19$). No youth-level demographic information was missing.

Normality Analyses

The mean, standard deviation, range, skewness, and kurtosis were examined for technical competence (CBAY-C model) and global (G-COMP) at the item and subscale level for the MMT and SMT conditions (see Tables 4, 5 and 6). Technical (CBAY-C model) and global (G-COMP)

items with a range of less than three were deemed problematic, due to restricted range (Jaccard & Becker, 2010). In the MMT condition, four technical competence (CBAY-C model) items had a range less than three: Emotion Education (n = 12, range = 2), Cognitive Anxiety (n = 3, range = 1.5), Coping Plan (n = 8, range = 2), Maintenance (n = 10, range = 2). In the MMT condition, the technical competence (CBAY-C model) subscale scores (M = 3.34, SD = .89) had a range of 4.33, and global competence (G-COMP) items had a range of at least 3.50. In the SMT condition, all competence items had a range of at least 3.50, technical competence (CBAY-C model) subscale scores (M = 3.86, SD = 1.01) had a range of 5.00, and global competence (G-COMP) items had a range of at least 4.50.

Skewness and kurtosis values are considered problematic if they fell outside the range of -2 to 2 (George & Mallery, 2016). In the MMT condition, technical competence (CBAY-C model) item skewness ranged from -.09 to 1.73, global competence (G-COMP) item skewness ranged from .24 to .54, and the technical competence (CBAY-C model) subscale skewness was .35. In the MMT condition, kurtosis values for technical competence (CBAY-C model) items ranged from -1.19 to 3.68, global competence (G-COMP) kurtosis values ranged from -.39 to 1.42, and the technical competence (CBAY-C model) subscale kurtosis was -.58. The "Exposure" item was the only item in the MMT condition deemed problematic because of its kurtosis value of 3.68. However, in the SMT condition, technical competence (CBAY-C model) item skewness ranged from -.14 to 1.01, global competence (G-COMP) item skewness ranged from .10 to .43, and the technical competence (CBAY-C model) subscale skewness was .36. Additionally, technical competence (CBAY-C model) item kurtosis ranged from -1.23 to 1.29, global competence (G-COMP) item kurtosis ranged from -.50 to .09, and the technical

competence (CBAY-C model) subscale kurtosis was -.26. Hence, all relevant competence item and subscale scores were normally distributed in the SMT condition.

Reliability and Subscale Generation: MMT Condition

Subscales for technical competence were generated from CBAY-C model items that were matched to each condition (i.e., MMT or SMT; see Data Analysis Plan for further detail). As in previous research (e.g., McLeod et al., 2018), scores generated by the two CBAY-C coders were averaged for each item for each case. Competence scores were averaged if both coders provided scores for an item. Next, CBAY-C model item scores were averaged to create CBAY-C model subscale scores. Similar procedures were used to create adherence (CBAY-A model) subscale scores; scores generated by the two CBAY-A coders were averaged for each item for each case. Then, CBAY-A model item scores were averaged to create model subscale scores.

Inter-rater reliability was also considered when generating technical competence subscales used in analyses; items that demonstrated "poor" (Cicchetti, 1994) inter-rater reliability were considered for exclusion. Per hypothesis 1, it was expected that technical competence (CBAY-C model) items would demonstrate "fair" to "excellent" (Cicchetti, 1994) inter-rater reliability, with a "good" (Cicchetti, 1994) mean ICC(2,2). In the MMT condition, CBAY-C model item inter-rater reliability ranged from ICC(2,2) = .17 to .89 (*M* = .53, *SD* = .23; see Table 4). The ICCs(2,2) for two of the nine CBAY-C model items fell within the "excellent" range, one item fell within the "good" range, three fell within the "fair" range, and three items (Coping Plan [*n* = 8, 95% CI (-2.32 – .87), ICC(2,2) = .34], Exposure [*n* = 61, 95% CI (-.02 — .46), ICC(2,2) = .38], and Maintenance [*n* = 10, 95% CI (-2.34 – .79)], ICC(2,2) = .17) fell within the "poor" range (Cicchetti, 1994).

Moreover, the CBAY-C model item "Maintenance" was excluded from the CBAY-C model subscale in the MMT condition because the inter-rater reliability for this item was "poor" (ICC[2,2] = .17; Cicchetti, 1994). (Note: CBAY-A model item "Maintenance" was also excluded from the CBAY-A model subscale to keep the content consistent across competence and adherence scales.) Similarly, the CBAY-C model items "Coping Plan" and "Exposure" were considered for exclusion from the CBAY-C model subscale in the MMT condition because these items also demonstrated "poor" inter-rater reliability (ICC[2,2] = .34 and .38, respectively; Cicchetti, 1994). However, these items were retained, as these practice elements are core components of CBT for youth anxiety.

Furthermore, G-COMP items were used to measure global competence. As in previous research (e.g., Brown et al., 2018), scores between the two G-COMP coders were averaged for each item for each case. Correlations between items of the G-COMP were examined to assess redundancy (see Tables 7 and 8). Although several G-COMP items were correlated above r = .70, no items were combined to follow Brown et al. (2018)'s suggestion of using the five G-COMP items as global competence scores.

Similarly, inter-rater reliability was considered when generating the global competence subscales (i.e., Alliance Building, Positive Expectancies, Focusing Treatment, Instigating Change, and Responsiveness). Hypothesis 2 stated that global competence (G-COMP) items would demonstrate "fair" to "excellent" (Cicchetti, 1994) inter-rater reliability, with a mean ICC(2,2) in the "good" (Cicchetti, 1994) range. In the MMT condition, G-COMP item inter-rater reliability ranged from ICC[2,2] = .49 to .59 (M = .53, SD = .04; see Table 6). All G-COMP items were retained, as these items all demonstrated at least "fair" (Cicchetti, 1994) inter-rater reliability.

Reliability and Subscale Generation: SMT Condition

The same procedures used in the MMT condition for matching technical competence (CBAY-C model) item content to the treatment program's manual were used in the SMT to identify applicable technical competence (CBAY-C model) items (see Data Analysis Plan for further detail). Scores generated by the two CBAY-C coders were averaged for each item for each case to create CBAY-C model scores; scores were averaged only if both coders provided scores for an item. Next, CBAY-C model item scores were averaged to create CBAY-C model subscale scores. Identical procedures were used to create adherence (CBAY-A model) subscale scores.

Inter-rater reliability was also considered when generating technical competence (CBAY-C model) subscales used in analyses; items that demonstrated "poor" (Cicchetti, 1994) inter-rater reliability were considered for exclusion. Hypothesis 1 stated that technical competence (CBAY-C model) items would demonstrate "fair" to "excellent" (Cicchetti, 1994) inter-rater reliability, with a "good" (Cicchetti, 1994) mean ICC(2,2). In the SMT condition, CBAY-C model item inter-rater reliability ranged from ICC(2,2) = .54 to .84 (*M* = .71, *SD* = .08; see Table 5). ICCs(2,2) for four of the 11 CBAY-C model items fell within the "excellent" range, six items fell within the "good" range, one fell within the "fair" range, and zero fell within the "poor" range (Cicchetti, 1994).

Per Hypothesis 2, it was expected that global competence (G-COMP) items (i.e., Alliance Building, Positive Expectancies, Focusing Treatment, Instigating Change, and Responsiveness) would demonstrate "fair" to "excellent" (Cicchetti, 1994) inter-rater reliability, with a mean ICC(2,2) in the "good" (Cicchetti, 1994) range. In the SMT condition, G-COMP item inter-rater reliability ranged from ICC(2,2) = .63 to .75 (*M* = .69, *SD* = .05; see Table 6). All competence

(CBAY-C model and G-COMP) items were retained in the SMT condition, as these items all demonstrated at least "fair" inter-rater reliability with "good" mean ICC(2,2)s (Cicchetti, 1994).

Construct Validity in the MMT Condition

To address the study's primary aim, technical competence (CBAY-C model), global competence (G-COMP), adherence (CBAY-A model), and alliance (TPOCS-A) scores were correlated with one another to determine whether technical and global competence could be measured distinctly in the MMT condition (Hypothesis 3; see Table 7 and Figures 1-5). Correlations between CBAY-C MMT and G-COMP subscale scores ranged from r = .41 to .80, with a mean of r = .59 (SD = .16). One correlation between CBAY-C MMT subscale and G-COMP: Instigating Change scores was r = .80, suggesting that these scores might be redundant (Kline, 1979). Overall, these findings suggested that CBAY-C MMT scores were distinguishable from G-COMP scores, except for the G-COMP subscale of Instigating Change.

Correlations between CBAY-C MMT, CBAY-A MMT, and TPOCS-A subscale scores were also examined (see Table 7). The correlation between CBAY-C MMT and CBAY-A MMT subscale scores was r = .11, whereas the correlation between CBAY-C MMT and TPOCS-A subscale scores was r = .07. These correlations were "small" in magnitude (Rosenthal & Rosnow, 1984). The mean of the correlations between CBAY-C MMT and G-COMP subscale scores (r = .59) was significantly higher than the correlation between CBAY-C MMT and CBAY-A MMT subscale scores (r = .11; z = 5.09; p < .001) and the correlation between CBAY-C MMT and TPOCS-A subscale scores (r = .07; z = 5.33; p <.001). That the CBAY-C MMT scores were more strongly correlated with the G-COMP scores than they were with the CBAY-A MMT or TPOCS-A scores supported the hypothesized pattern. Additionally, that CBAY-C MMT scores were more strongly correlated to CBAY-A MMT subscale scores than they were to

TPOCS-A subscale scores (albeit, not at a statistically significant level, $z = .35$; $p = .36$) partially supported the hypothesized pattern.

Correlations between G-COMP, CBAY-A MMT, and TPOCS-A subscale scores were also examined (see Table 7). Correlations between G-COMP and CBAY-A MMT subscale scores ranged from $r = .06$ to .30 and averaged $r = .14$ ($SD = .10$), whereas correlations between G-COMP and TPOCS-A subscale scores ranged from $r = .20$ to .43 and averaged $r = .32$ ($SD = .10$). The mean inter-item correlation between G-COMP and CBAY-A MMT subscale scores was "small" in magnitude (Rosenthal & Rosnow, 1984), and the mean inter-item correlation between G-COMP and TPOCS-A subscale scores was "medium" (Rosenthal & Rosnow, 1984). The mean inter-item correlation between CBAY-C MMT and G-COMP subscale scores ($r = .59$) was significantly higher than the mean inter-item correlation between G-COMP and CBAY-A MMT subscale scores ($r = .14$; $z = 5.27$; $p < .001$) and the mean inter-item correlation between G-COMP and TPOCS-A subscale scores ($r = .32$; $z = 3.32$; $p < .001$). Thus, G-COMP scores were more strongly correlated with CBAY-C MMT scores than they were with CBAY-A MMT or TPOCS-A scores, which supported the hypothesized pattern. Yet, the mean inter-item correlation between G-COMP and CBAY-A MMT subscale scores was significantly smaller in magnitude than the mean inter-item correlation between G-COMP and TPOCS-A subscale scores ($z = -2.04$; $p = .02$), which did not support the hypothesized pattern.

Because these findings suggested some redundancy (Kline, 1979) between CBAY-C MMT and G-COMP subscale scores, a post-hoc principal components analysis was conducted to examine whether the two domains of competence were distinct in the MMT condition (see Table 9). The CBAY-C MMT subscale and the five G-COMP item scores were analyzed using principal component analysis with Oblique (non-orthogonal) rotation. Components were

extracted if eigenvalues were greater than one (Kaiser, 1960), and the analysis yielded two components explaining a total of 81.8% of the variance for the entire set of variables. Component one explained 63.9% of the variance and included G-COMP: Positive Expectancies, G-COMP: Focusing Treatment, G-COMP: Instigating Change, and the CBAY-C model subscale. Whereas, Component 2 explained 17.9% of the variance and included G-COMP: Alliance Building and G-COMP: Responsiveness. These findings suggested that there may be two components that described the competence items in the MMT condition, but they did not describe technical and global competence distinctly.

Construct Validity in the SMT Condition

Technical competence (CBAY-C model), global competence (G-COMP), adherence (CBAY-A model), and alliance (TPOCS-A) scores were correlated to determine whether technical and global competence could be measured distinctly in the SMT condition (Hypothesis 3; see Table 8 and Figures 6-10). Correlations between CBAY-C SMT and G-COMP subscale scores ranged from r = .65 to .85, with a mean of r = .73 (SD = .09). The correlation between the CBAY-C SMT subscale and G-COMP: Focusing Treatment scores was r = .79, suggesting that these scores might be redundant (Kline, 1979). Additionally, the correlation between CBAY-C SMT subscale and G-COMP: Instigating Change scores was r = .85, suggesting that these scores might also be redundant (Kline, 1979). Overall, these findings suggested that CBAY-C SMT scores were not distinguishable from G-COMP scores.

Correlations between CBAY-C SMT, CBAY-A SMT, and TPOCS-A subscale scores were also examined (see Table 8). The correlation between CBAY-C SMT and CBAY-A SMT subscale scores was r = .24, whereas the correlation between CBAY-C SMT and TPOCS-A subscale scores was r = .44. The correlation between CBAY-C SMT and CBAY-A SMT

subscale scores was “medium” in magnitude (Rosenthal & Rosnow, 1984), and the correlation between CBAY-C SMT and TPOCS-A subscale scores was “large” (Rosenthal & Rosnow, 1984). The mean inter-item correlation between CBAY-C SMT and G-COMP subscale scores (r = .73) was significantly higher than the correlation between CBAY-C SMT and CBAY-A SMT subscale scores (r = .24; z = 8.4; p < .001) and the correlation between CBAY-C SMT and TPOCS-A subscale scores (r = .44; z = 5.53; p < .001). That CBAY-C SMT subscale scores were more strongly correlated with G-COMP item scores than they were with CBAY-A SMT or TPOCS-A subscale scores supported the hypothesized pattern. However, the correlation between CBAY-C SMT and CBAY-A SMT subscale scores was significantly smaller than the correlation between CBAY-C SMT and TPOCS-A subscale scores (z = -2.75; p < .001), which did not support the hypothesized pattern.

Correlations between G-COMP, CBAY-A SMT, and TPOCS-A subscale scores were also examined (see Table 8). Correlations between G-COMP and CBAY-A SMT subscale scores ranged from r = .05 to .42 and averaged r = .21 (SD = .16), whereas correlations between G-COMP and TPOCS-A subscale scores ranged from r = .34 to .63 and averaged r = .51 (SD = .12). The mean inter-item correlation between G-COMP and CBAY-A SMT subscale scores was “medium”, and the mean inter-item correlation between G-COMP and TPOCS-A subscale scores was “large” (Rosenthal & Rosnow, 1984). The mean inter-item correlation between CBAY-C SMT and G-COMP subscale scores (r = .73) was significantly higher than the mean inter-item correlation between G-COMP and CBAY-A SMT subscale scores (r = .21; z = 9.15; p < .001) and the mean inter-item correlation between G-COMP and TPOCS-A subscale scores (r = .51; z = 4.59; p < .001). That the G-COMP subscale scores were more strongly correlated with CBAY-C SMT subscale scores than they were with CBAY-A SMT or TPOCS-A subscale scores

supported the hypothesized pattern. However, the mean inter-item correlation between G-COMP and CBAY-A SMT subscale scores was significantly smaller in magnitude than the mean inter-item correlation between G-COMP and TPOCS-A subscale scores (z = -4.57; p < .001), which did not support the hypothesized pattern.

Because these findings suggested some redundancy (Kline, 1979) between CBAY-C SMT and G-COMP subscale scores, a post-hoc principal components analysis was conducted to examine whether the two domains of competence were distinct in the SMT condition (see Table 10). The CBAY-C SMT subscale and the five G-COMP subscale scores were analyzed using principal component analysis with Oblique (non-orthogonal) rotation. Components were extracted if eigenvalues were greater than one (Kaiser, 1960), and the analysis yielded one component explaining a total of 77% of the variance for the entire set of variables. These findings suggested that there was only one component that described the competence items in the SMT condition, and that technical and global competence were not measured distinctly.

Discussion

To date, the youth treatment literature does not provide a clear answer as to whether technical and global competence can be measured separately. This is important to determine, as these two domains of competence have been distinguished theoretically and may have implications for the measurement of competence as well as therapist training (Brown et al., 2018; Sburlati et al., 2011). The purpose of the current study was to determine whether technical and global competence could be measured as distinct constructs in CBT programs for youth anxiety. In the current study, there was insufficient evidence to demonstrate that technical competence could be consistently measured as distinct from all dimensions of global competence across two different CBT programs for youth anxiety. That is, while correlations between technical and

global competence subscale scores were mostly non-redundant (Kline, 1979), there were a few correlations between technical competence and certain global competence subscale scores (e.g., Instigating Change) that were considered redundant (Kline, 1979). This indicated that only some aspects of global competence were able to be measured distinctly from technical competence. Additionally, reliability of competence items and patterns of correlations amongst competence, adherence, and alliance did fully reflect findings of previous research (e.g., Brown et al., 2018; Hogue et al., 2008; McLeod et al., 2018). Lastly, while overall conclusions were drawn from both study conditions, individual findings across conditions were not identical.

Prior to examining the primary study aims, inter-rater reliability was assessed for all competence items. Both technical and global competence scores demonstrated higher inter-rater reliability in the standard manualized treatment (SMT) condition than in the modular manualized treatment (MMT) condition. A few technical competence items in the MMT condition demonstrated "poor" reliability (i.e., Coping Plan, Exposure, Maintenance; Cicchetti, 1994), whereas the remainder of the competence items demonstrated at least "fair" (Cicchetti, 1994) inter-rater reliability. In the SMT condition, all competence items demonstrated at least "fair" (Cicchetti, 1994) inter-rater reliability, with a "good" (Cicchetti, 1994) mean ICC(2,2). Thus, inter-rater reliability in the SMT condition mirrored previously demonstrated ICC(2,2)s of these competence scores. For instance, McLeod et al. (2018) found that the CBAY-C model items demonstrated "poor" to "excellent" (Cicchetti, 1994) inter-rater reliability at the item-level, with a "good" (Cicchetti, 1994) mean ICC(2,2) in a standard manualized CBT program for youth anxiety (i.e., Coping Cat). Similarly, Brown et al. (2018) found that G-COMP items demonstrated "good" to "excellent" (Cicchetti, 1994) inter-rater reliability, with a "good" (Cicchetti, 1994) mean ICC(2,2) in a sample of Coping Cat sessions. Thus, the competence items

in the SMT condition demonstrated similar inter-rater reliability as these competence items have in the past when used in the context of standard manualized CBT programs for youth anxiety. However, these previous studies (i.e., Brown et al., 2018; McLeod et al., 2018) did not examine the inter-rater reliability of CBAY-C or G-COMP scores in the context of modular manualized treatment sessions. Therefore, there was no direct comparison to be drawn between any previous study's use of the CBAY-C or G-COMP in the context of modular manualized CBT programs for youth anxiety.

While inter-rater reliability of competence items in the MMT condition might not have mirrored the findings of Brown et al. (2018) or McLeod et al. (2018), they were similar to the findings of other previous studies. For instance, Bjaastad et al. (2016) found that competence items demonstrated "fair" to "good" (Cicchetti, 1994) inter-rater reliability, with a "fair" (Cicchetti, 1994) mean ICC(2,2) in individual and group CBT programs for youth anxiety. Moreover, Hogue et al. (2008) found that technical competence items demonstrated mostly "poor" (Cicchetti, 1994) inter-rater reliability, with a "poor" (Cicchetti, 1994) mean ICC(2,2) in the context of a modular CBT program for youth substance use. Hence, although not as high as hypothesized, the "fair" (Cicchetti, 1994) mean inter-rater reliability of technical competence scores in the MMT condition was still greater than those previously found in other youth CBT programs (Bjaastad et al., 2016; Hogue et al., 2008). Thus, issues of "poor" (Cicchetti, 1994) inter-rater reliability at the item-level were not deemed a hindrance when patterns of correlations were considered in subsequent analyses conducted at the subscale level.

After examining inter-rater reliability, the study's primary hypothesis was addressed (i.e., whether technical and global competence could be measured as distinct constructs in CBT programs for youth anxiety). In the MMT condition, the mean inter-item correlation between

technical and global competence scores was non-redundant (Kline, 1979), which suggested that perhaps technical and global competence were measured as distinct constructs. Conversely, in the SMT condition, the mean inter-item correlation between technical and global competence scores was redundant (Kline, 1979), which suggested that perhaps technical and global competence were not measured as distinct constructs.

Besides mean inter-item correlations, individual correlations at the subscale level were also examined. Most correlations between technical and global competence subscale scores were non-redundant (Kline, 1979) across the two conditions. However, there were two global competence subscales (i.e., Focusing Treatment and Instigating Change) that were strongly correlated with technical competence in both conditions. In the MMT condition, Focusing Treatment was strongly correlated with technical competence, but Instigating Change was the only global competence subscale score considered redundant (Kline, 1979) with technical competence scores. Similarly, both Focusing Treatment and Instigating Change were strongly correlated with technical competence scores in the SMT condition at a level that suggested redundancy (Kline, 1979). Thus, correlations amongst technical and global competence subscale scores suggested that perhaps some aspects of global competence (i.e., alliance building, positive expectancies, responsiveness), but not all dimensions of global competence (i.e., focusing treatment and instigating change), could be measured distinctly from technical competence in CBT programs for youth anxiety. This finding is similar to what Brown et al. (2018) demonstrated when they correlated CBAY-C and G-COMP scores and found that Focusing Treatment and Instigating Change were the dimensions of global competence most strongly associated with technical competence.

The structured nature of CBT is one possible explanation for why these two dimensions of global competence (i.e., focusing treatment and instigating change) were so closely associated with technical competence in the current sample. Specifically, focusing clients on treatment tasks as well as encouraging clients to participate in treatment activities (e.g., encouraging clients to participate in exposures) are important elements of CBT for youth anxiety (e.g., Hofmann et al., 2012; Kazdin & Weisz, 1998). Thus, as focusing treatment and instigating change are somewhat inherent in the implementation of CBT practice elements, it follows that the global competence subscale scores of Focusing Treatment and Instigating Change would be so strongly associated with technical competence scores in the current study's CBT for youth anxiety sample.

Lastly, principal component analysis findings were also assessed to examine whether technical and global competence could be measured distinctly in the current study. Neither condition's principal component analysis results suggested that technical and global competence were measured as distinct constructs. More specifically, the principal component analysis conducted with the SMT condition's data suggested that only one component was captured by the technical and global competence subscale scores. Additionally, while there were two components that were extracted from the analysis conducted with the MMT condition's data, the two components did not separate into technical and global competence. Instead, technical competence and three global competence items (i.e., Positive Expectancies, Focusing Treatment, and Instigating Change) loaded onto one component; whereas, two global competence items (Alliance Building and Responsiveness) loaded onto the second component. Hence, the findings of the principal component analyses suggested that perhaps the competence subscales in the current study did not measure technical and global competence distinctly. Taken together, the current study's findings did not provide enough evidence to suggest that all dimensions of global

competence could be measured distinctly from technical competence across CBT programs for youth anxiety.

The mixed findings of the current study are similar to the varied findings from previous literature. Some researchers have found that technical and global competence could be measured distinctly from one another. For example, Hogue et al. (2008) found a "large" (Rosenthal & Rosnow, 1984) yet non-redundant (Kline, 1979) correlation between technical and global competence scores in a CBT program for youth substance use (r = .68). Whereas, other researchers found that technical and global competence could not be measured as distinct constructs. For instance, Bjaastad et al. (2016) found that technical and global competence scores were strongly correlated (r = .88) at a level that suggested redundancy (Kline, 1979) in CBT programs for youth anxiety. McLeod et al. (2018) also found "large" (Rosenthal & Rosnow, 1984) mean inter-item correlations between technical competence items and overall therapist skillfulness (M = .77, SD = .12) and responsiveness (M = .80, SD = .08) at a level that suggested redundancy (Kline, 1979) in a standard manualized CBT program for youth anxiety. However, none of these studies examined the distinction between technical and global competence by using the two instruments utilized in the current study. Because these previous studies did not operationalize technical or global competence as precisely as the current study, it was also important to consider studies that had conceptualized competence similarly to the current study.

Notably, the current study's findings did not fully mirror the findings of Brown et al. (2018)'s study, which was the most similar to the current study with regards to competence instruments (i.e., both studies used CBAY-C and G-COMP) and sample (i.e., both studies' samples were youth receiving CBT for anxiety). Brown et al. (2018) found a "large" (Rosenthal & Rosnow, 1984) yet non-redundant (Kline, 1979) mean inter-item correlation between technical

and global competence scores (M = .36, SD = .15). Whereas, in the current study, correlations between technical and global competence scores were larger in magnitude than those found in Brown et al. (2018)'s study (z = -2.43, p = .008 in the MMT condition; z = -7.94, p < .001 in the SMT condition). Some of the correlations between technical and global competence scores in the current study were also so large in magnitude that they were considered redundant (Kline, 1979). Thus, technical and global competence were more easily distinguished in Brown et al. (2018)'s study than in the current study. Perhaps one explanation of the different findings between the two similar studies is that Brown et al. (2018) utilized different teams of coders to rate technical (CBAY-C) and global (G-COMP) competence; whereas, the current study utilized the same team of coders to rate both rate technical (CBAY-C) and global (G-COMP) competence. The use of two different teams of coders for technical and global competence could have added to the distinction between the two competence domains (Brown et al., 2018; Campbell & Fiske, 1959). Altogether, the current study's findings provided insufficient evidence to suggest technical and global competence could be measured distinctly in CBT programs for youth anxiety.

Furthermore, the construct validity of technical and global competence scores was also examined in the current study by comparing competence scores to adherence and alliance scores. In the current study, mean inter-item correlations between technical and global competence were "large" (Rosenthal & Rosnow, 1984) and greater in magnitude than the correlations between competence and adherence scores and between competence and alliance scores in both study conditions. This pattern is similar to what has been demonstrated in previous research. For instance, McLeod et al. (2018) found "large" (Rosenthal & Rosnow, 1984) mean inter-item correlations between scores of technical and global competence (r's = .77 to .80) that were greater in magnitude than mean inter-item correlations between competence and adherence and

between competence and alliance scores in a standard manualized CBT program. Hogue et al. (2008) also found a "large" (Rosenthal & Rosnow, 1984) correlation between technical and global competence scores (r = .68) that was greater in magnitude than correlations between competence and adherence scores and between competence and alliance scores in a CBT program for youth substance use. Thus, the discriminant validity of competence scores was supported in the current study, such that the competence scores were distinct from both adherence and alliance scores. The convergent validity of competence scores was also supported in the current study, such that technical and global competence scores were strongly correlated with one each other.

Yet, competence scores were more strongly associated with alliance scores than with adherence scores across conditions in the current study, which did not support the hypothesized pattern. Specifically, correlations between competence and adherence (r's = .21 to .24) were smaller in magnitude than the correlations between competence and alliance (r's = .44 to .51) in the SMT condition. Correlations between competence and adherence (r's = .11 to .14) were also smaller in magnitude than the correlations between competence and alliance (r's = .07 to .32) in the MMT condition. The only exception to this pattern was that the correlation between technical competence and adherence (r = .11) was larger than the correlation between technical competence and alliance (r = .07) in the MMT condition. However, the difference between these two correlations was not statistically significant; thus, competence scores were deemed to be more strongly associated with alliance scores than adherence scores across conditions in the current study.

The pattern found in the current study was unlike what was found in previous research, as previous studies have generally found that competence scores were more strongly correlated to

adherence scores than to alliance scores. For instance, McLeod et al. (2018) found a "large" (Rosenthal & Rosnow, 1984) mean inter-item correlation between technical competence and adherence scores (r = .43) and a "small" (Rosenthal & Rosnow, 1984) mean inter-item correlation between technical competence and alliance scores (r = .21) in a standard manualized CBT program. Similarly, Brown et al. (2018) also found a "large" (Rosenthal & Rosnow, 1984) mean inter-item correlation between global competence and adherence scores (r = .39) and found "small" to "medium" (Rosenthal & Rosnow, 1984) mean inter-item correlations between global competence and alliance scores (r's = .22 to .32) in a standard manualized CBT program. Likewise, Hogue et al. (2008) found "large" (Rosenthal & Rosnow, 1984) correlations between competence and adherence scores (r's = .42 to .50) and a "medium" correlation between competence and alliance scores (r = .31) in a CBT condition.

While the pattern between competence, adherence, and alliance scores found in the current study did not exactly mirror the pattern found in previous research (i.e., Brown et al., 2018; Hogue et al., 2008; McLeod et al., 2018), it may be reflective of previously posited theories. That is, some researchers have theorized that competence and alliance are closely related with regards to the effectiveness of treatment. For instance, competence has been conceptualized as a mechanism that strengthens the effectiveness of treatment by bolstering the alliance between therapist and client (Smith, Dishion, Shaw, & Wilson, 2013). Moreover, alliance building has been categorized as a component of global competence (e.g., Brown et al., 2018). That competence scores were more strongly associated with alliance than adherence scores in the current study perhaps provides evidence to support certain conceptualizations of competence (e.g., Smith et al., 2013) over previous empirical findings (e.g., Brown et al., 2018; Hogue et al., 2008; McLeod et al., 2018).

Another secondary conclusion drawn from the present findings was that the competence instruments seemed to perform differently across the two conditions. For instance, there were several instances of "poor" (Cicchetti, 1994) inter-rater reliability in the MMT condition but not in the SMT condition. It could have been that certain technical competence items (e.g., Exposure) were more difficult to code in the MMT condition than in the SMT condition because of the more unstructured and adaptive nature of treatment in the MMT condition. Moreover, both competence instruments (CBAY-C and G-COMP) were developed using samples of treatment sessions of standard manualized CBT for youth anxiety (i.e. Coping Cat; Brown et al., 2018; McLeod et al., 2018). Thus, perhaps the wording of certain items was biased towards detection in the SMT condition versus the MMT condition. The different performance of the competence items (i.e., differences in inter-rater reliability) across conditions might suggest that the competence items used in the current study could be slightly more suitable for standard manualized CBT programs as opposed to modular manualized CBT programs.

Moreover, there were differences in validity analysis findings across the two conditions as well. For instance, the mean inter-item correlation between technical and global competence scores was non-redundant (Kline, 1979) in the MMT condition but was redundant (Kline, 1979) in the SMT condition. Similarly, the correlation between G-COMP: Focusing Treatment and technical competence subscale scores was non-redundant (Kline, 1979) in the MMT condition but was redundant (Kline, 1979) in the SMT condition. Lastly, competence scores were all more strongly correlated with alliance scores than adherence scores in the SMT condition. Yet, technical competence scores were equally associated with both adherence and alliance scores in the MMT condition. These findings also suggested that the competence instruments used in the current study performed slightly differently across the two study conditions.

There were also several limitations of the current study that should be noted. One such limitation was the limited sample of treatment programs used in the current study. Because the current study only examined technical and global competence in youth CBT programs for anxiety, conclusions cannot be generalized to other treatment modalities. For instance, it is unknown if technical and global competence scores would be more or less distinct in a client-centered therapy approach. Similarly, since CBT is a particularly structured and focused treatment program (e.g., Hofmann et al., 2012; Kazdin & Weisz, 1998), it may have been easier to detect examples of technical competence (i.e., quality of the implementation of specific practice elements) than it was to differentially detect global competence (i.e., individual instances of a therapist being responsive outside of the implementation of practice elements). Lastly, another limitation of the current study was that the same coders rated both competence instruments. In the current study, it could be that some of the redundancy in competence scores was due to the same coder rating both technical (CBAY-C) and global (G-COMP) competence (Brown et al., 2018; Campbell & Fiske, 1959).

Given the mixed results of the current study and aforementioned limitations, there are several potential avenues of future research that should be explored. One possible future study would be to examine the ability of coders to use the CBAY-C in another sample of MATCH sessions. Perhaps there was something unique to the current sample that made it difficult for coders to achieve acceptable inter-rater reliability on all CBAY-C items in the MMT condition. Conversely, perhaps the CBAY-C is not as suitable for use with modular manualized CBT programs as it is for standard manualized CBT programs. Additionally, because technical competence instruments are specific to the types of treatment in which they are utilized, it would be helpful to test the current study's hypotheses in the context of different treatment modalities

(e.g., dialectical behavioral therapy, client-centered) for other problems (e.g., depression, oppositional defiant disorder). Perhaps technical and global competence cannot be fully distinguished in CBT for youth anxiety programs but would be able to be distinguished in other treatment programs. Lastly, future research should continue to examine how competence is conceptualized and assessed. Another possible explanation for the overlap between technical and global competence in the current study is that "responsiveness" is a key feature of both components of competence. Perhaps skillfulness should be used to define technical competence, whereas responsiveness could be reserved for defining global competence. Future studies might also examine the predictive validity of technical and global competence instruments to further test whether these two constructs are only theoretically distinct and not empirically or practically distinct.

In sum, the current study's findings did not provide enough evidence to conclude that technical and global competence could be consistently measured as distinct constructs across CBT programs for youth anxiety. Instead, the current findings suggested that perhaps only certain aspects of global competence (i.e., alliance building, positive expectancies, and responsiveness) could be measured distinctly from technical competence in CBT for youth anxiety. Therefore, as previous research (e.g., Bjaastad et al., 2016; Hogue et al., 2008; McLeod et al., 2018) suggested, there still remains insufficient evidence to conclude that technical and global competence can or should be measured as different constructs in CBT programs for youth anxiety.

References

Achenbach TM, Rescorla LA. *Manual for the ASEBA school-age forms & profiles*. Burlington: Research Centre for Children, Youth and Families, University of Vermont; 2001.

American Psychological Association (2002). Ethical principles of psychologists and code of conduct. *American Psychologist, 57,* 1060-1073.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*. Washington, DC: American Psychiatric Association.

Barber, J. P., Sharpless, B. A., Klostermann, S., & McCarthy, K. S. (2007). Assessing intervention competence and its relation to therapy outcome: A selected review derived from the outcome literature. *Professional Psychology: Research and Practice*, *38*(5), 493.

Barkley, R. A. (2013). *Defiant children: A clinician's manual for assessment and parent training*. Guilford press.

Bjaastad, J. F., Haugland, B. S. M., Fjermestad, K. W., Torsheim, T., Havik, O. E., Heiervang, E. R., & Öst, L. G. (2016). Competence and Adherence Scale for Cognitive Behavioral Therapy (CAS-CBT) for anxiety disorders in youth: Psychometric properties. *Psychological Assessment*, *28*(8), 908.

Brown, R. C., Southam-Gerow, M. A., McLeod, B. D., Wheat, E. B., Tully, C. B., Reise, S. P., ... & Weisz, J. R. (2018). The global therapist competence scale for youth psychosocial treatment: Development and initial validation. *Journal of Clinical Psychology*, *74*(4), 649-664.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81.

Campbell, M. K., Elbourne, D. R., & Altman, D. G. (2004). CONSORT statement: Extension to

cluster randomised trials. *BMJ*, *328*(7441), 702-708.

Carroll, K. M., Nich, C., Sifry, R. L., Nuro, K. F., Frankforter, T. L., Ball, S. A., ... & Rounsaville, B. J. (2000). A general system for evaluating therapist adherence and competence in psychotherapy research in the addictions. *Drug and Alcohol Dependence*, *57*(3), 225-238.

Castonguay, L. G. (1993). " Common factors" and" nonspecific variables": Clarification of the two concepts and recommendations for research. *Journal of Psychotherapy Integration*, *3*(3), 267.

Castonguay, L. G., & Beutler, L. E. (2006). Principles of therapeutic change: A task force on participants, relationships, and techniques factors. *Journal of clinical psychology*, *62*(6), 631-638.

Chorpita, B. F., & Daleiden, E. L. (2009). Mapping evidence-based treatments for children and adolescents: Application of the distillation and matching model to 615 treatments from 322 randomized trials. *Journal of Consulting and Clinical Psychology*, *77*(3), 566–579.

Chorpita, B. F., & Weisz, J. R. (2005). *MATCH-ADC: A Modular Approach to Treatment for Children with Anxiety, Depression, and Conduct Problems*. Honolulu, HI, and Boston, MA: University of Hawaii at Manoa and Judge Baker Children's Center, Harvard Medical School.

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.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

DeVellis, R. F. (2017). *Scale development: Theory and applications*, 4th ed. Los Angeles, CA: Sage Publications.

Garber, J., Keiley, M. K., & Martin, N. C. (2002). Developmental trajectories of adolescents' depressive symptoms: predictors of change. *Journal of Consulting and Clinical Psychology*, *70*(1), 79.

Eyberg, S. M., Schuhmann, E. M., & Rey, J. (1998). Child and adolescent psychotherapy research: Developmental issues. *Journal of Abnormal Child Psychology*, *26*(1), 71-82.

Forgatch, M. S., Patterson, G. R., & DeGarmo, D. S. (2005). Evaluating fidelity: Predictive validity for a measure of competent adherence to the Oregon Model of Parent Management Training. *Behavior Therapy*, *36*(1), 3.

Foster, S. L., & Cone, J. D. (1995). Validity issues in clinical assessment. *Psychological Assessment*, *7*(3), 248.

Fouad, N. A., Grus, C. L., Hatcher, R. L., Kaslow, N. J., Hutchings, P. S., Madson, M. B., ... & Crossman, R. E. (2009). Competency benchmarks: A model for understanding and measuring competence in professional psychology across training levels. *Training and Education in Professional Psychology*, *3*(4S), S5.

Frank, J. D. (1971). Therapeutic factors in psychotherapy. *American Journal of Psychotherapy*, *25*, 350–361.

George, D., & Mallery, P. (2016). Descriptive statistics. In *IBM SPSS Statistics 23 Step by Step* (pp. 126-134). Routledge.

Gutermann, J., Schreiber, F., Matulis, S., Stangier, U., Rosner, R., & Steil, R. (2015).

Therapeutic adherence and competence scales for Developmentally Adapted Cognitive Processing Therapy for adolescents with PTSD. *European Journal of Psychotraumatology*, *6*(1), 26632.

Hofmann, S. G., Asnaani, A., Vonk, I. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research*, *36*(5), 427-440.

Hogue, A., Dauber, S., Stambaugh, L. F., Cecero, J. J., & Liddle, H. A. (2006). Early therapeutic alliance and treatment outcome in individual and family therapy for adolescent behavior problems. *Journal of Consulting and Clinical Psychology*, *74*(1), 121.

Hogue, A., Dauber, S., Chinchilla, P., Fried, A., Henderson, C., Inclan, J., ... & Liddle, H. A. (2008). Assessing fidelity in individual and family therapy for adolescent substance abuse. *Journal of Substance Abuse Treatment*, *35*(2), 137-147.

Humphreys, L., Crino, R., & Wilson, I. (2018). The competencies movement: Origins, limitations, and future directions. *Clinical Psychologist*, *22*(3), 290-299.

Jaccard, J., & Becker, M. A. (2010). *Statistics for the Behavioral Sciences* (5th ed.). Belmont, CA: Wadsworth.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*(1), 141-151.

Kaslow, N. J. (2004). Competencies in professional psychology. *American Psychologist*, *59*(8), 774.

Kazdin, A. E. (2016). *Research Design in Clinical Psychology* (5th ed.). Boston, MA: Pearson.

Kazdin, A. E., & Weisz, J. R. (1998). Identifying and developing empirically supported child and adolescent treatments. *Journal of Consulting and Clinical Psychology*, *66*(1), 19.

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

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.

Kline, P. (1979). Psychometrics and psychology. London, UK: Academic Press.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155-163.

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

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

Margolin, G., Oliver, P., Gordis, E., O'Hearn, H., Medina, A., Ghosh, C., & Morland, L. (1998). The nuts and bolts of behavioral observation of marital and family interaction. *Clinical Child & Family Psychology Review, 1*, 195–213.

Marvin, S. E., Miklowitz, D. J., O'brien, M. P., & Cannon, T. D. (2016). Family-focused therapy for individuals at clinical high risk for psychosis: treatment fidelity within a multisite randomized trial. *Early Intervention in Psychiatry, 10*(2), 137-143.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica: Biochemia Medica, 22*(3), 276-282.

McLeod, B. D., Southam-Gerow, M. A., Rodríguez, A., Quinoy, A. M., Arnold, C. C., Kendall, P. C., & Weisz, J. R. (2018). Development and initial psychometrics for a therapist competence instrument for CBT for youth anxiety. *Journal of Clinical Child & Adolescent Psychology*, *47*(1), 47-60.

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.

Norcross, J. C. (2002). *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients*. Oxford University Press.

Norcross, J. C., & Wampold, B. E. (2011). Evidence-based therapy relationships: Research conclusions and clinical practices. *Psychotherapy*, *48*(1), 98.

Orlinsky, D. E., & Howard, K. I. (1987). A generic model of psychotherapy. *Journal of Integrative & Eclectic Psychotherapy*.

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

Podell, J. L., Kendall, P. C., Gosch, E. A., Compton, S. N., March, J. S., Albano, A. M., ... & Keeton, C. P. (2013). Therapist factors and outcomes in CBT for anxiety in youth. *Professional Psychology: Research and Practice*, *44*(2), 89.

Ramsawh, H. J., Chavira, D. A., & Stein, M. B. (2010). Burden of anxiety disorders in pediatric medical settings: prevalence, phenomenology, and a research agenda. *Archives of Pediatrics & Adolescent Medicine*, *164*(10), 965-72.

Resko, S. M., Walton, M. A., Chermack, S. T., Blow, F. C., & Cunningham, R. M. (2012).

Therapist competence and treatment adherence for a brief intervention addressing alcohol and violence among adolescents. *Journal of Substance Abuse Treatment, 42*(4), 429-437.

Rosenthal, R., & Rosnow, R. L. (1984). Essentials of behavioral research: Methods and data analysis. New York, NY: McGraw-Hill.

Sburlati, E. S., Schniering, C. A., Lyneham, H. J., & Rapee, R. M. (2011). A model of therapist competencies for the empirically supported cognitive behavioral treatment of child and adolescent anxiety and depressive disorders. *Clinical Child and Family Psychology Review*, *14*(1), 89-109.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147-177.

Sharpless, B. A., & Barber, J. P. (2009). A conceptual and empirical review of the meaning, measurement, development, and teaching of intervention competence in clinical psychology. *Clinical Psychology Review*, *29*(1), 47-56.

Shirk, S. R., & Russell, R. L. (1998). Process issues in child psychotherapy. In A. Bellack & M. Hersen (Eds.), *Comprehensive Clinical Psychology, Vol. 15*, pp. 57-82. Oxford: Pergamon

Shirk, S. R., & Saiz, C. (1992). The therapeutic alliance in child therapy: Clinical, empirical, and developmental perspectives. *Development and Psychopathology, 4*(4), 713-728.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420-428.

Smith, J. D., Dishion, T. J., Shaw, D. S., & Wilson, M. N. (2013). Indirect effects of fidelity to the family check-up on changes in parenting and early childhood problem behaviors. *Journal of Consulting and Clinical Psychology*, 81(6), 962.

Southam-Gerow, M. A., McLeod, B. D., Arnold, C. A., Rodriguez, A., Cox, J. R., Reise, S. P. … Kendall, P. C. (2016). Initial development of a treatment adherence measure for cognitive-behavioral therapy for child anxiety. *Psychological Assessment*, *28*(1), 70–80.

Spruill, J., Rozensky, R. H., Stigall, T. T., Vasquez, M., Bingham, R. P., & Olvey, C. D. V. (2004). Becoming a competent clinician: Basic competencies in intervention. *Journal of Clinical Psychology*, *60*(7), 741-754.

Stallard, P., Myles, P., & Branson, A. (2014). The cognitive behaviour therapy scale for children and young people (CBTS-CYP): Development and psychometric properties. *Behavioural and Cognitive Psychotherapy*, *42*(3), 269-282.

Tabachnik, B. G., & Fidell, S. L. (2007). *Using multivariate statistics*. Boston: Pearson Education Inc.

Weck, F., Hilling, C., Schermelleh-Engel, K., Rudari, V., & Stangier, U. (2011). Reliability of adherence and competence assessment in cognitive behavioral therapy: Influence of clinical experience. *The Journal of Nervous and Mental Disease, 199*(4), 276-279.

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

Weisz, J. R., Weersing, V. R., Valeri, S. M., & McCarty, C. A. (1999). *Therapist's Manual PASCET: Primary and secondary control enhancement training program*. Los Angeles: University of California.

Weisz, J. R., Chorpita, B. F., Frye, A., Ng, M. Y., Lau, N., Bearman, S. K., ... & Hoagwood, K.

E. (2011). Youth Top Problems: Using idiographic, consumer-guided assessment to identify treatment needs and to track change during psychotherapy. *Journal of Consulting and Clinical Psychology, 79*(3), 369-380.

Weller, E. B., Weller, R. A., Fristad, M. A., Rooney, M. T., & Schecter, J. (2000). Children's interview for psychiatric syndromes (ChIPS). *Journal of the American Academy of Child & Adolescent Psychiatry, 39*(1), 76-84.

Tables and Figures

Table 1. *CBAY-C subscales for MMT (MATCH) and SMT (Coping Cat)*

Modular Manualized Treatment (MMT)	Standard Manualized Treatment (SMT)
Model: 1. Psychoeducation 2. Emotion Education 3. Fear Ladder 4. Cognitive Anxiety 5. Coping Plan 6. Exposure: Prep 7. Exposure 8. Exposure: Debrief 9. Maintenance	Model: 1. Psychoeducation 2. Emotion Education 3. Fear Ladder 4. Relaxation 5. Cognitive Anxiety 6. Problem Solving 7. Self-Reward 8. Coping Plan 9. Exposure: Prep 10. Exposure 11. Exposure: Debrief

Table 2. *Youth and Therapist Descriptive Data and Group Comparisons: Excluded vs. Included*

Variable	**M (SD) or %**		
Youth	**MMT/SMT** (N = 38)	**Excluded** (N = 6)	***t* or χ^2 value (p-value)**
Age	9.84 (1.65)	11.00 (2.35)	1.41 (.17)
Sex			
Female	47.37	16.67	1.99 (.16)
Race/Ethnicity			5.73 (.33)
Caucasian	60.53	16.67	
African American	5.26	.00	
Asian American	2.63	16.67	
Latinx	2.63	.00	
Multiracial	26.32	33.34	
Other	2.63	.00	
Not Reported	.00	33.34	
CBCL (pre) scores			
Total	64.58 (8.73)	60.17 (10.27)	-1.13 (.27)
Internalizing	69.82 (7.81)	65.67 (11.27)	-1.14 (.26)
Externalizing	57.34 (11.45)	50.83 (11.20)	-1.30 (.20)
Anxiety	69.74 (7.34)	63.67 (10.33)	-1.78 (.08)
Therapist	**MMT/SMT** (N = 26)		
Age	40.34 (9.67)		
Sex			
Female	80.8		
Race/Ethnicity			
Caucasian	53.8		
African American	7.7		
Asian American	23.1		
Other	7.7		
Not Reported	7.7		
Years of Experience	6.79 (8.09)		
Area of Specialty			
Social worker	38.5		
Behavior specialist	19.2		

Psychologist Mental health counselor Not reported Other	15.4 19.2 3.8 3.8		

Table 3. *Youth and Therapist Descriptive Data and Group Comparisons: MMT vs. SMT*

Variable	**M (SD) or %**		
Youth	**MMT** (N = 16)	**SMT** (N = 22)	***t* or χ^2 value (p-value)**
Age	9.94 (1.88)	9.77 (1.51)	.30 (.77)
Sex			.15 (.70)
Female	43.75	50.00	
Race/Ethnicity			8.18 (.15)
Caucasian	43.75	72.72	
African American	12.50	.00	
Asian American	.00	4.54	
Latinx	.00	4.54	
Multiracial	37.50	18.18	
Other	6.25	.00	
CBCL (pre) scores			
Total	63.63 (10.39)	65.27 (7.49)	-.57 (.57)
Internalizing	69.56 (9.33)	70.00 (6.72)	-.17 (.87)
Externalizing	55.06 (11.64)	59.00 (11.28)	-1.05 (.30)
Anxiety	69.88 (7.76)	69.64 (7.20)	.10 (.92)
Therapist	**MMT** (N = 10)	**SMT** (N = 16)	***t* or χ^2 value (p-value)**
Age	35.20 (6.81)	43.56 (9.96)	-2.33 (.03)
Sex			.01 (.94)
Female	80.00	81.25	
Years of Experience	3.67 (1.68)	8.67 (9.80)	-1.51 (.15)
Race/Ethnicity			4.67 (.32)
Caucasian	50.00	56.25	
African American	.00	12.50	
Asian American	40.00	12.50	
Not reported	.00	12.50	
Other	10.00	6.25	
Area of Specialty			3.19 (.67)
Social worker	40.00	37.50	
Behavior specialist	30.00	12.50	
Psychologist	20.00	12.50	

Mental health counselor	10.00	25.00	
Not reported	.00	6.25	
Other	.00	6.25	

Table 4. *CBAY-C Item and Subscale Descriptive Data and Inter-rater Reliability: MMT*

Item Type	Item	*N*	Range	Minimum	Maximum	*M*	*SD*	Skewness	Kurtosis	ICC(2,2)
Model	Psychoeducation	67	4.00	2.00	6.00	3.68	.75	.27	.18	.48
Model	Emotion Education	12	2.00	2.50	4.50	3.50	.64	.31	-.86	.44
Model	Fear Ladder	34	4.00	1.50	5.50	3.66	1.05	.02	-.84	.83
Model	Cognitive Anxiety	3	1.50	3.00	4.50	3.50	.87	1.73		.89
Model	Coping Plan	8	2.00	3.00	5.00	3.94	.73	-.09	-1.19	.34
Model	Exposure: Preparation	64	3.50	1.50	5.00	2.65	.70	.92	.85	.57
Model	Exposure	61	4.00	2.00	6.00	2.97	.84	1.7	3.68	.38
Model	Exposure: Debrief	44	3.00	2.00	5.00	3.02	.83	.71	-.22	.63
Model	Maintenance	10	2.00	3.00	5.00	4.00	.62	.00	-.91	.17
	Model Subscale	164	4.33	1.67	6.00	3.34	.89	.35	-.58	.68

Note. *N* represents the number of times an item was rated as present at least once during a treatment session by either coder.

Table 5. *CBAY-C Item and Subscale Descriptive Data and Inter-rater Reliability: SMT*

Item Type	Item	N	Range	Minimum	Maximum	Mean	SD	Skewness	Kurtosis	ICC(2,2)
Model	Psychoeducation	55	3.50	2.50	6.00	3.70	.84	.51	.07	.67
Model	Emotion Education	65	5.00	2.00	7.00	4.31	1.03	.43	.47	.78
Model	Fear Ladder	50	4.00	2.00	6.00	3.66	1.07	.39	-.68	.76
Model	Relaxation	38	3.50	2.50	6.00	3.99	.82	.83	1.29	.54
Model	Cognitive Anxiety	33	4.50	2.00	6.50	4.41	1.00	-.14	-.12	.65
Model	Problem Solving	11	4.00	2.50	6.50	4.5	1.30	.00	-1.23	.67
Model	Self-Reward	14	3.50	3.00	6.50	4.18	1.12	.77	-.52	.84
Model	Coping Plan	91	4.50	2.00	6.50	4.03	.93	.45	-.13	.70
Model	Exposure: Preparation	68	5.00	1.50	6.50	3.35	1.23	.58	-.64	.80
Model	Exposure	59	4.50	2.00	6.50	3.39	1.00	1.01	.62	.67
Model	Exposure: Debrief	41	4.50	2.00	6.50	3.65	1.20	.42	-.72	.73
	Model Subscale	305	5.00	2.00	7.00	3.86	1.01	.36	-.26	.77

Note. N represents the number of times an item was rated as present at least once during a treatment session by either coder.

Table 6. *G-COMP Item Descriptive Data and Inter-rater Reliability*

Item	*N*	Range	Minimum	Maximum	*M*	*SD*	Skewness	Kurtosis	ICC(2,2)
MMT									
Alliance Building	244	5.00	1.50	6.50	4.08	.73	.36	1.42	.49
Positive Expectancies	244	3.50	2.00	5.50	3.69	.78	.24	-.39	.51
Focusing Treatment	244	4.00	2.00	6.00	3.40	.90	.54	-.32	.59
Instigating Change	244	4.00	1.50	5.50	3.34	.74	.44	-.30	.53
Responsiveness	244	4.00	2.00	6.00	3.76	.79	.38	.25	.51
SMT									
Alliance Building	359	5.00	2.00	7.00	4.44	.96	.10	.05	.68
Positive Expectancies	359	4.50	2.00	6.50	3.94	.93	.37	.05	.63
Focusing Treatment	359	4.50	2.00	6.50	3.87	1.00	.25	-.50	.64
Instigating Change	359	5.50	1.50	7.00	3.90	1.08	.43	-.27	.75
Responsiveness	359	5.50	1.50	7.00	4.21	1.11	.17	.09	.73

Note. N represents the number of times an item was rated as present at least once during a treatment session by either coder.

Table 7. *Correlations between Competence, Adherence, and Alliance Subscale Scores: MMT*

	1	2	3	4	5	6	7	8
1. G-COMP 1	1	.57**	.22**	.42**	.78**	.46**	.07	.43**
N	244	244	244	244	244	164	244	219
2. G-COMP 2		1	.55**	.68**	.52**	.66**	.13*	.38**
N		244	244	244	244	164	244	219
3. G-COMP 3			1	.75**	.22**	.64**	.30**	.22**
N				244	244	164	244	219
4. G-COMP 4				1	.45**	.80**	.15*	.20**
N				244	244	164	244	219
5. G-COMP 5					1	.41**	0.06	.38**
N					244	164	244	219
6. CBAY-C: Model						1	.11	.07
N						164	164	150
7. CBAY-A: Model							1	.26**
N							244	219
8. TPOCS-A								1
N								219

Note: G-COMP 1 = Alliance Building, G-COMP 2 = Positive Expectancies, G-COMP 3 = Focusing Treatment, G-COMP 4 = Instigating Change, G-COMP 5 = Responsiveness

Table 8. *Correlations between Competence, Adherence, and Alliance Subscale Scores: SMT*

	1	2	3	4	5	6	7	8
1. G-COMP 1	1	.78**	.48**	.76**	.89**	.67**	0.05	.60**
N	359	359	359	359	359	305	359	332
2. G-COMP 2		1	.58**	.78**	.74**	.65**	.16**	.48**
N		359	359	359	359	305	359	332
3. G-COMP 3			1	.78**	.49**	.79**	.42**	.36**
N			359	359	359	305	359	332
4. G-COMP 4				1	.78**	.85**	.32**	.49**
N				359	359	305	359	332
5. G-COMP 5					1	.67**	0.08	.63**
N					359	305	359	332
6. CBAY-C: Model						1	.24**	.44**
N						305	305	288
7. CBAY-A: Model							1	.14*
N							359	332
8. TPOCS-A								1
N								332

Note: G-COMP 1 = Alliance Building, G-COMP 2 = Positive Expectancies, G-COMP 3 = Focusing Treatment, G-COMP 4 = Instigating Change, G-COMP 5 = Responsiveness

Table 9. *Principal Component Analysis for MMT Competence Subscales*

	Loadings	
	Component 1	Component 2
G-COMP 1	.27	**.90**
G-COMP 2	**.68**	.50
G-COMP 3	**.88**	.07
G-COMP 4	**.87**	.33
G-COMP 5	.19	**.92**
CBAY-C MMT Subscale	**.85**	.27
Eigenvalue	3.84	1.07
% of Total Variance	63.94	17.87
Total Variance		**81.81%**

Note: G-COMP 1 = Alliance Building, G-COMP 2 = Positive Expectancies, G-COMP 3 = Focusing Treatment, G-COMP 4 = Instigating Change, G-COMP 5 = Responsiveness

Table 10. *Principal Component Analysis for SMT Competence Subscales*

	Loadings
	Component 1
G-COMP 1	**.89**
G-COMP 2	**.87**
G-COMP 3	**.79**
G-COMP 4	**.95**
G-COMP 5	**.88**
CBAY-C SMT	**.88**
Subscale	
Eigenvalue	4.62
% of Total Variance	77.01
Total Variance	77.01

Note: G-COMP 1 = Alliance Building, G-COMP 2 = Positive Expectancies, G-COMP 3 = Focusing Treatment, G-COMP 4 = Instigating Change, G-COMP 5 = Responsiveness

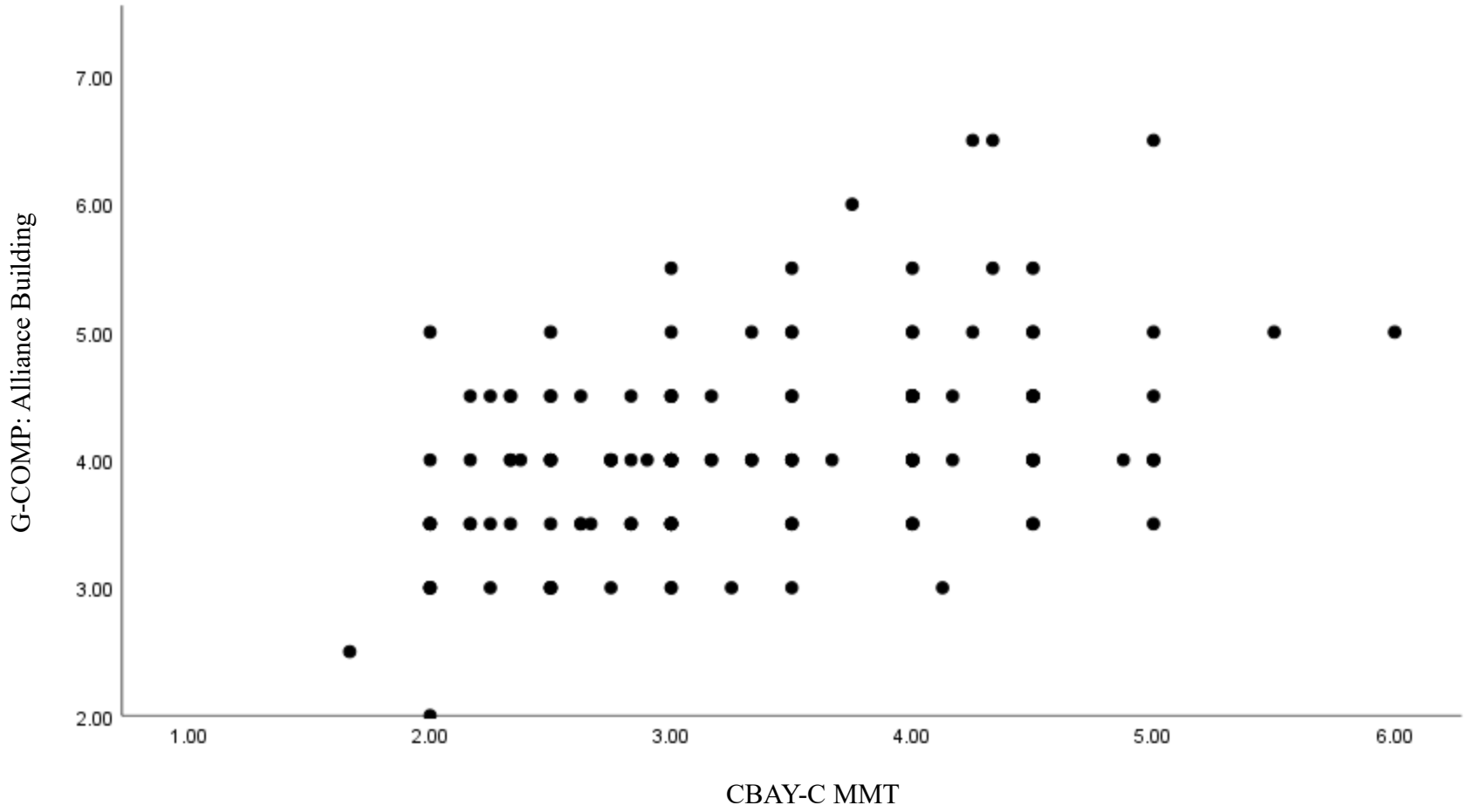

Figure 1. Correlation between CBAY-C MMT and G-COMP: Alliance Building subscale scores in the MMT condition.

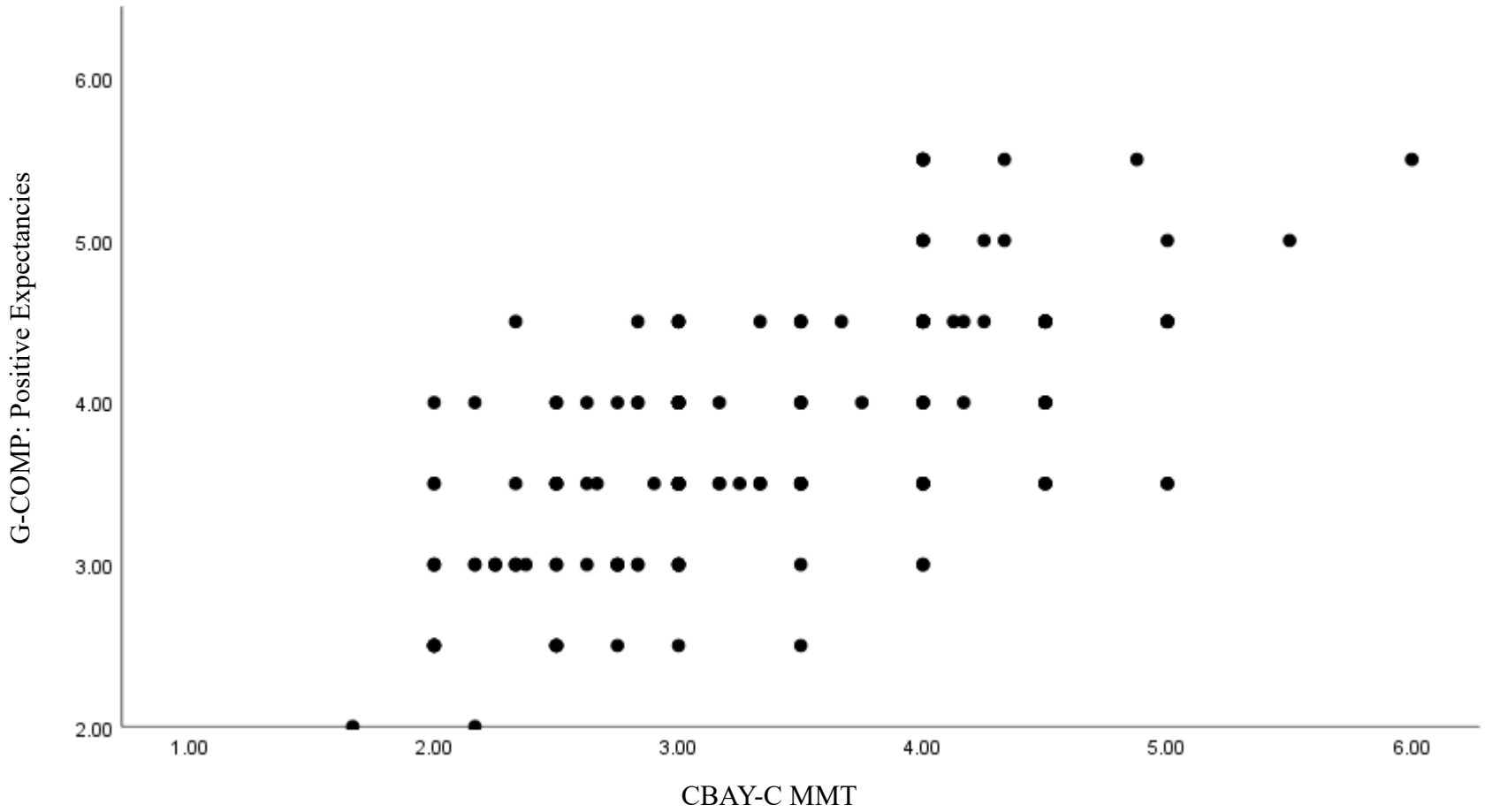

Figure 2. Correlation between CBAY-C MMT and G-COMP: Positive Expectancies subscale scores in the MMT condition.

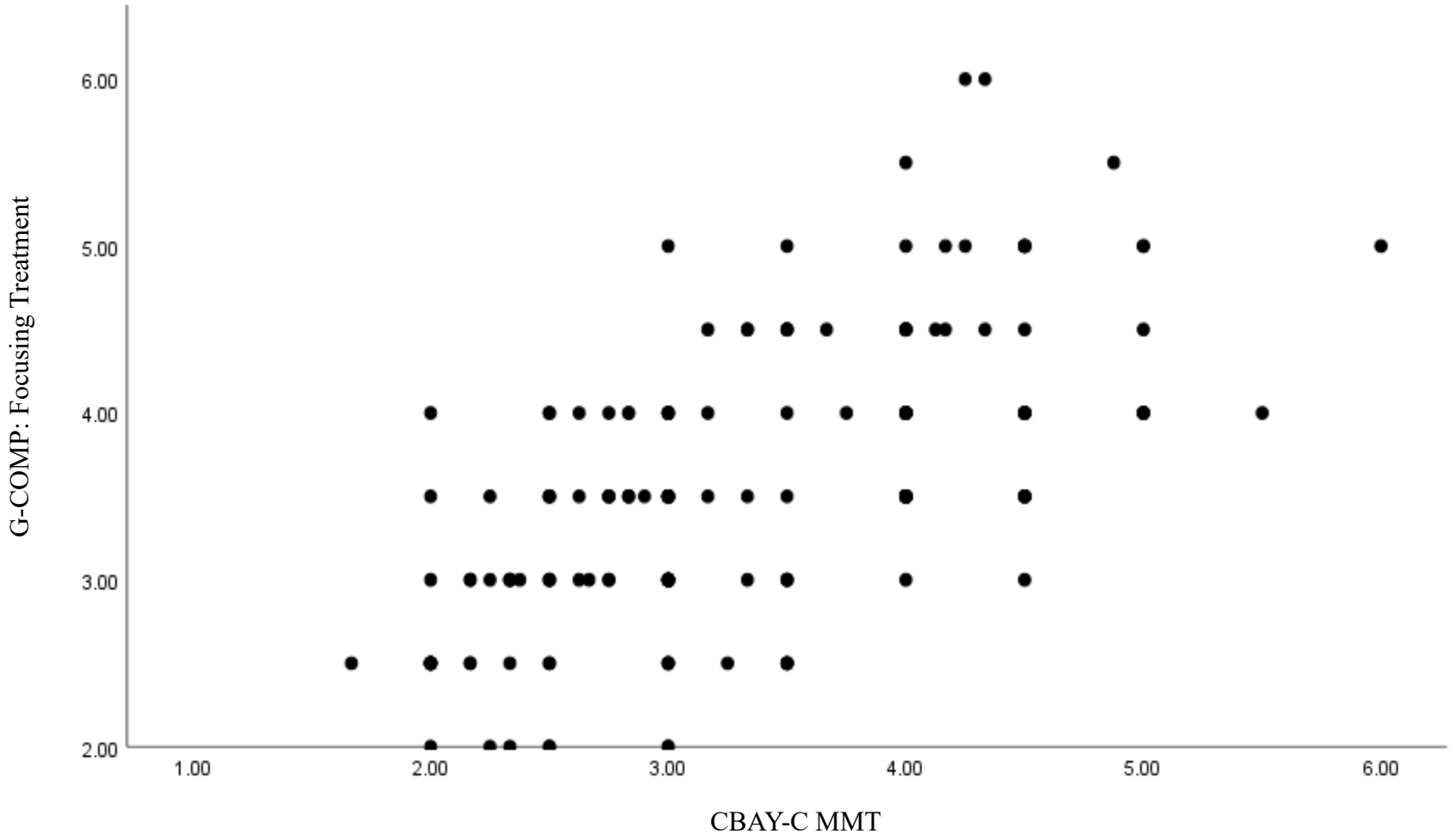

Figure 3. Correlation between CBAY-C MMT and G-COMP: Focusing Treatment subscale scores in the MMT condition.

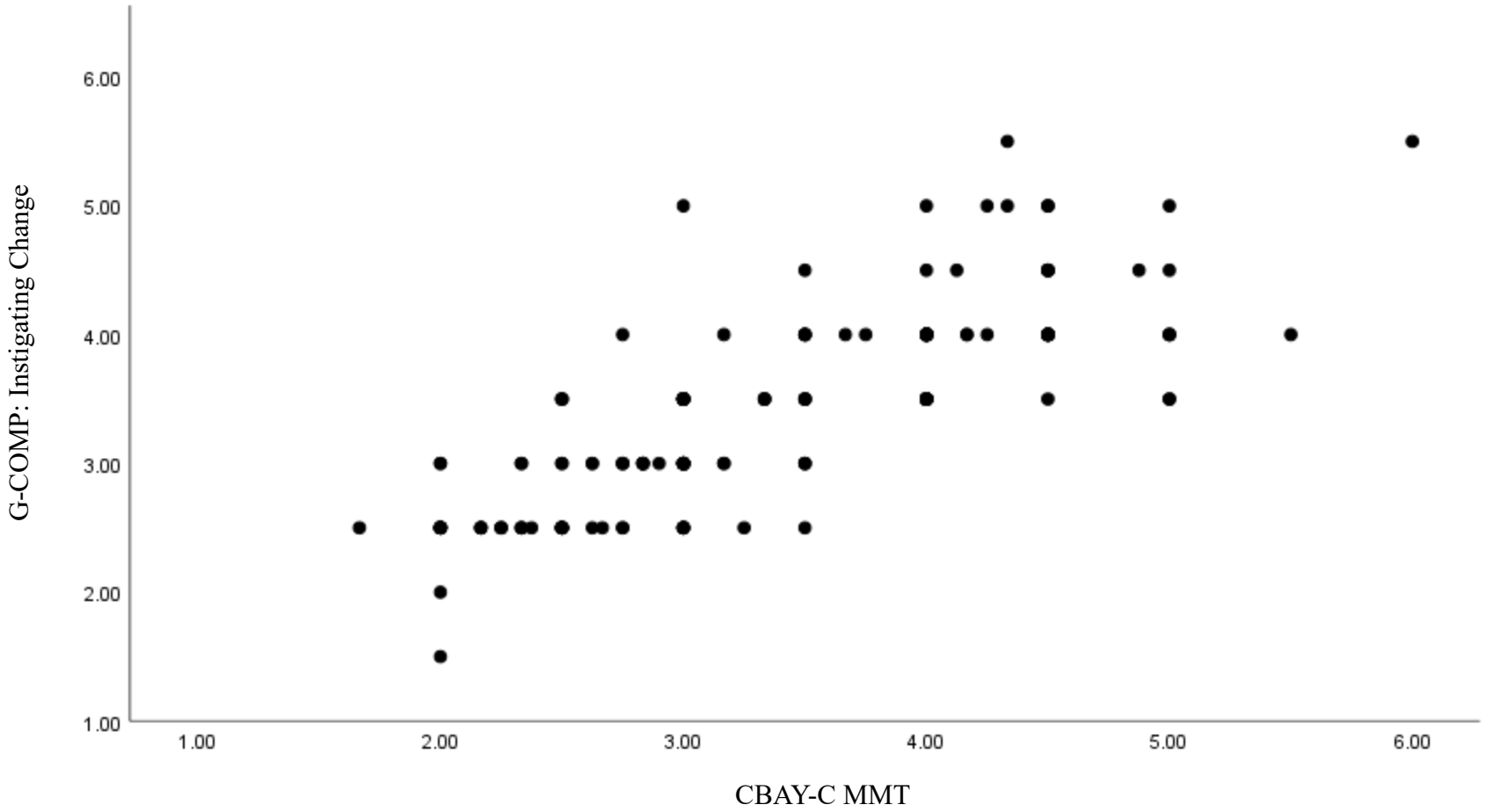

Figure 4. Correlation between CBAY-C MMT and G-COMP: Instigating Change subscale scores in the MMT condition.

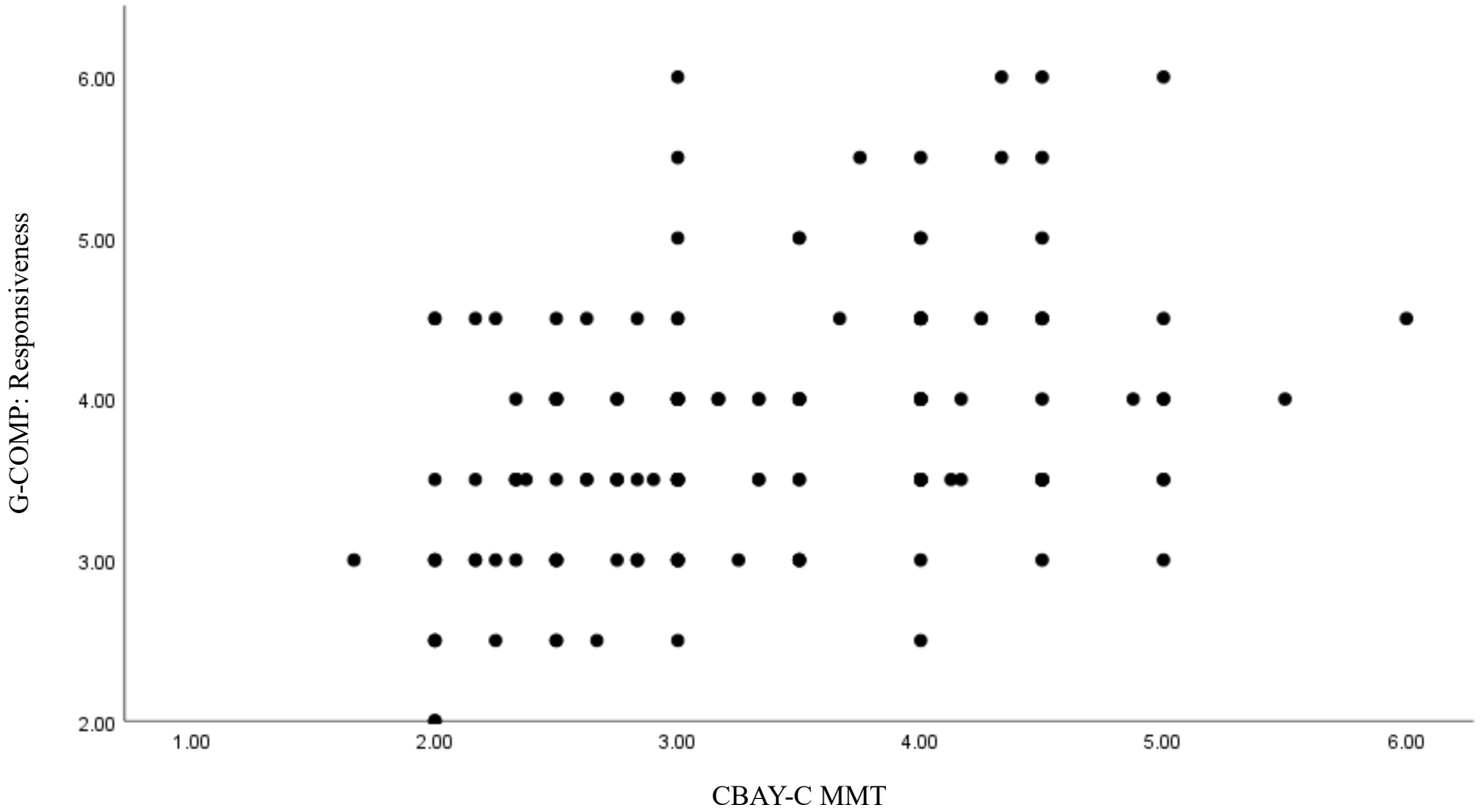

Figure 5. Correlation between CBAY-C MMT and G-COMP: Responsiveness subscale scores in the MMT condition.

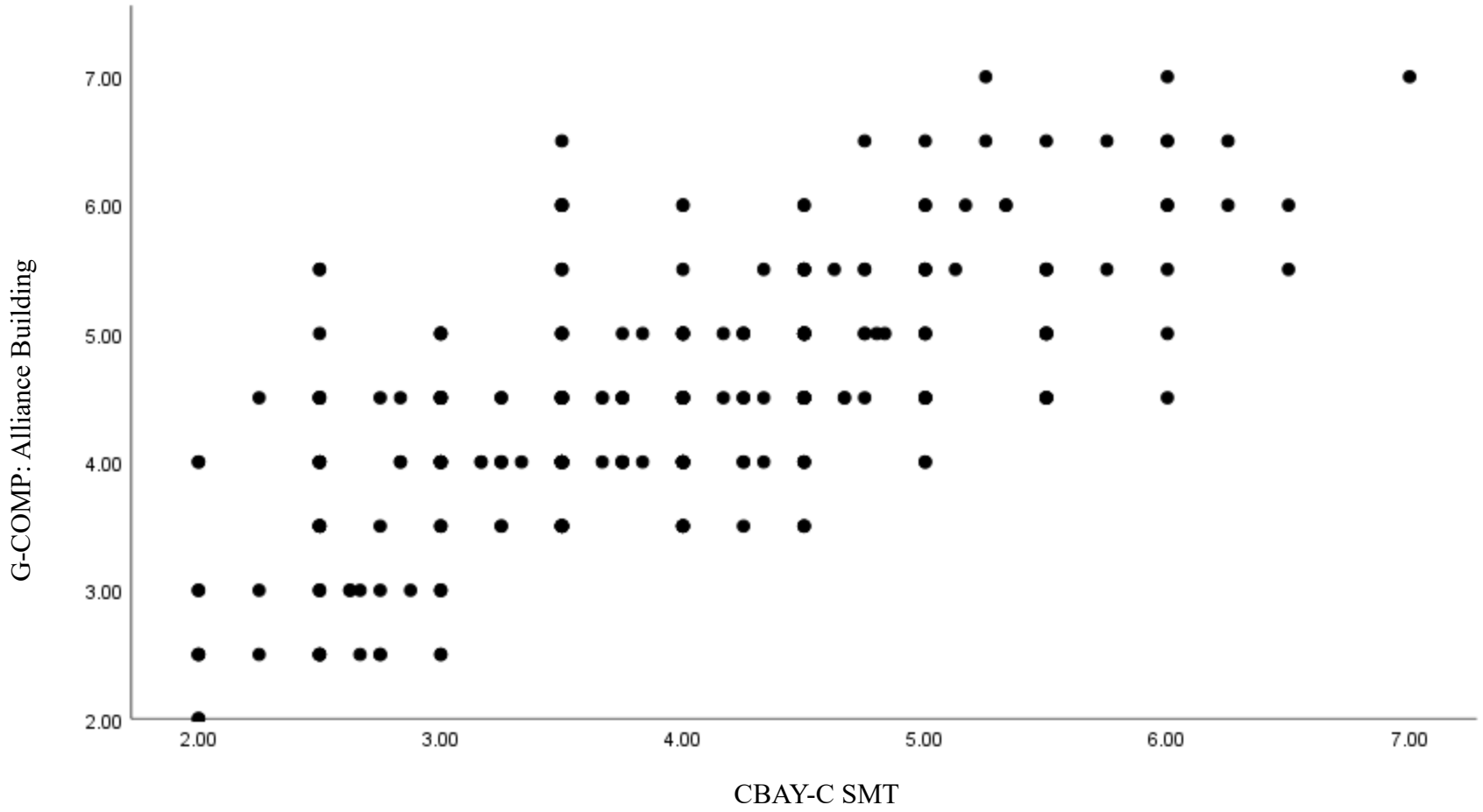

Figure 6. Correlation between CBAY-C SMT and G-COMP: Alliance Building subscale scores in the SMT condition.

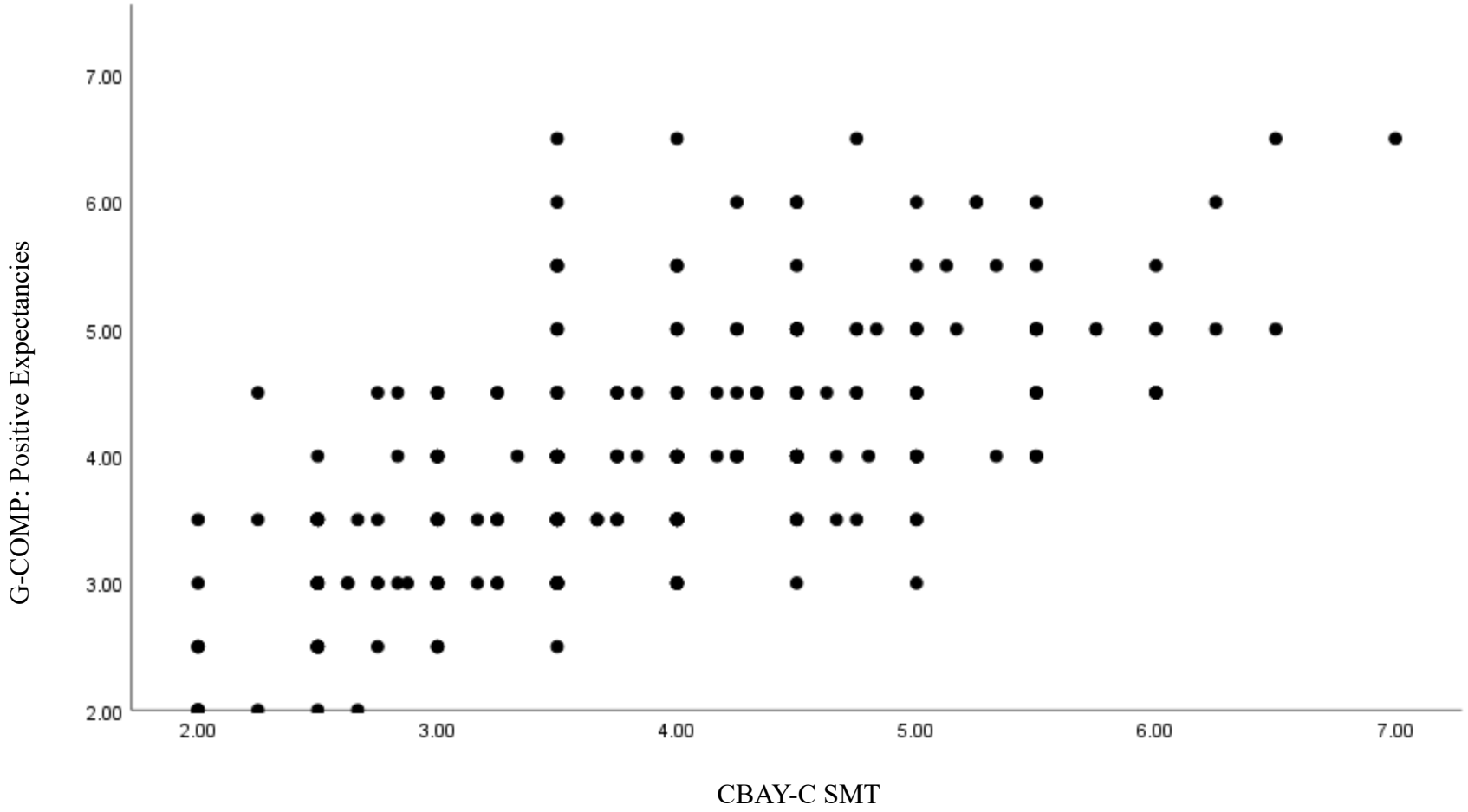

Figure 7. Correlation between CBAY-C SMT and G-COMP: Positive Expectancies subscale scores in the SMT condition.

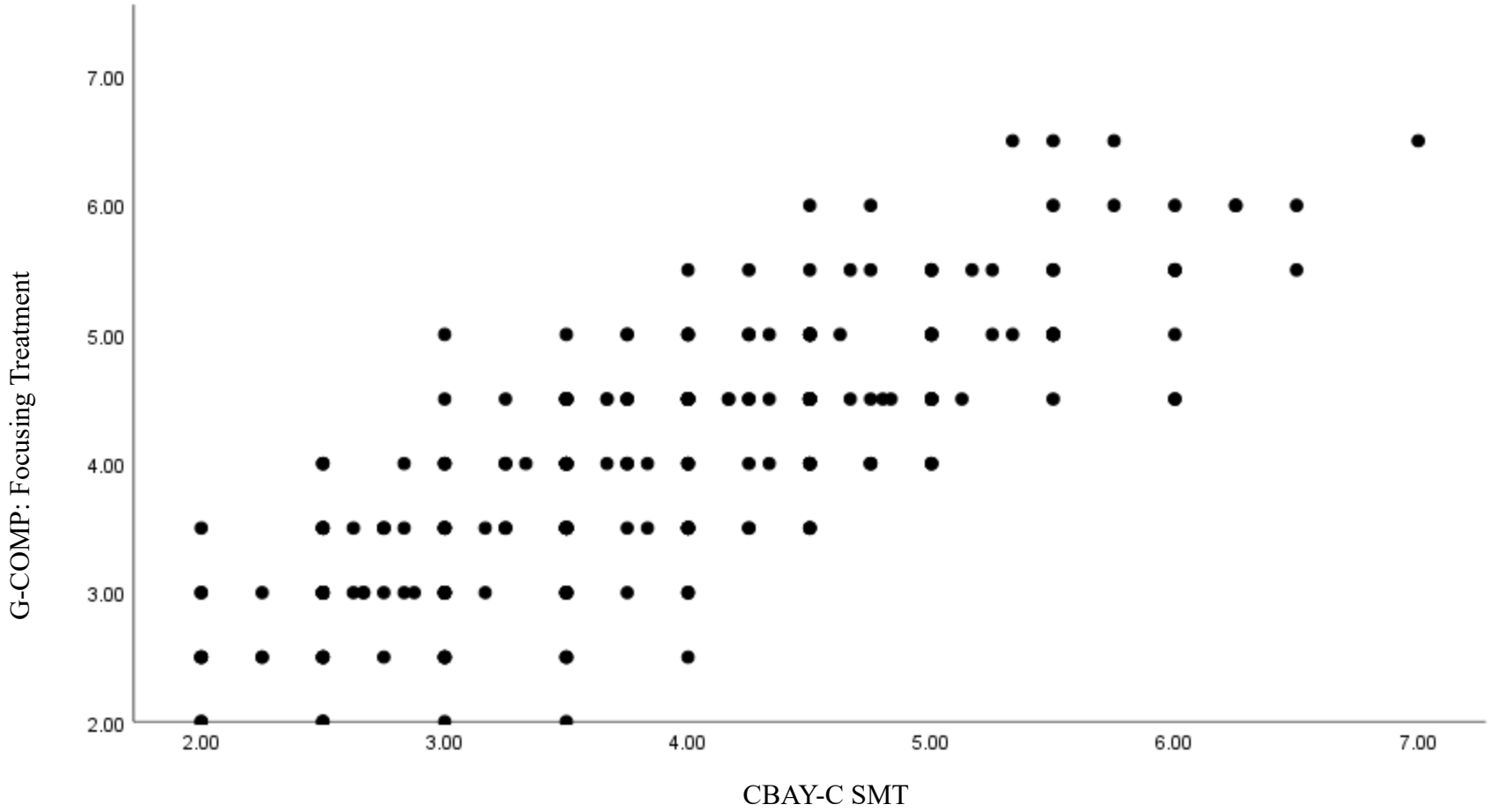

Figure 8. Correlation between CBAY-C SMT and G-COMP: Focusing Treatment subscale scores in the SMT condition.

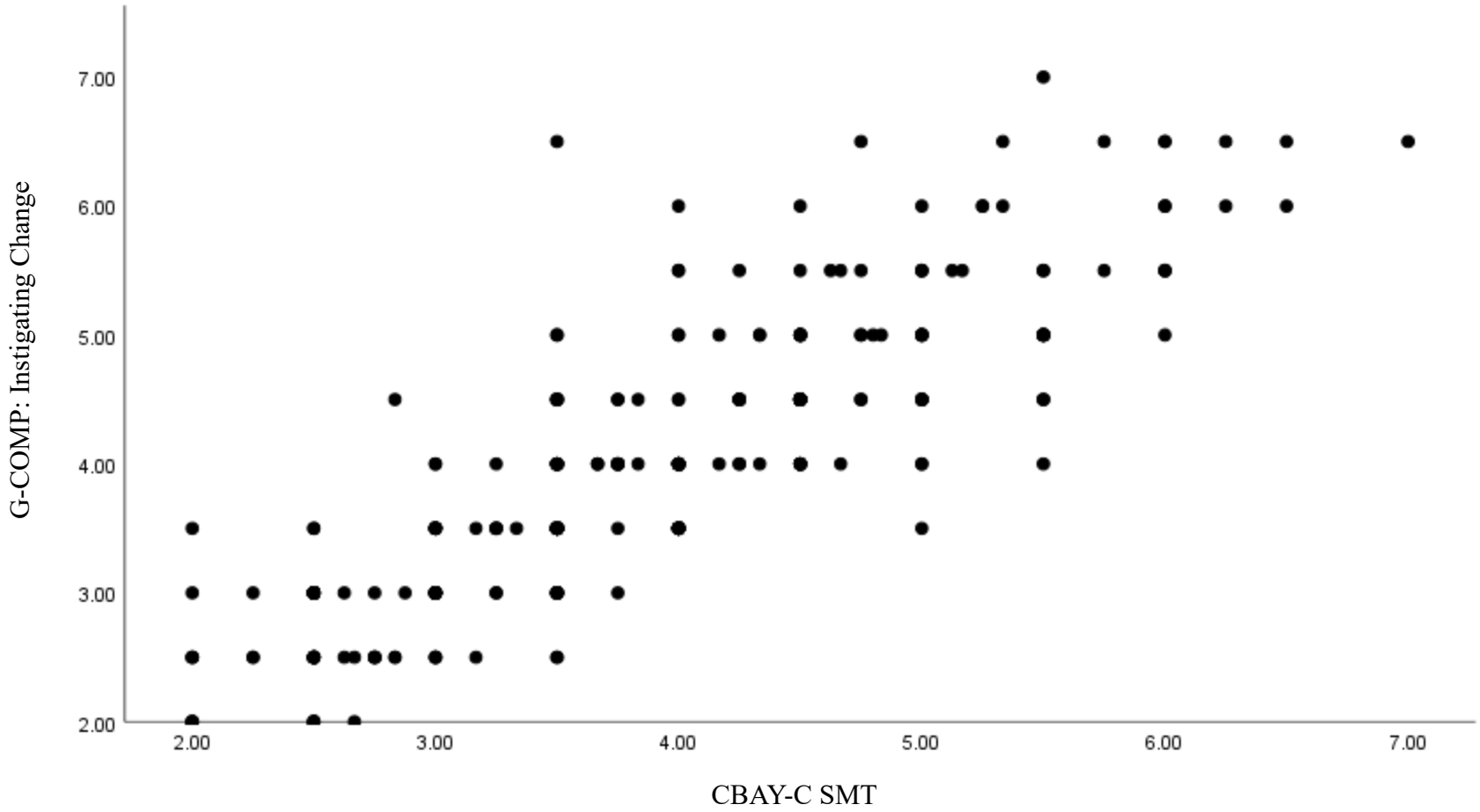

Figure 9. Correlation between CBAY-C SMT and G-COMP: Instigating Change subscale scores in the SMT condition.

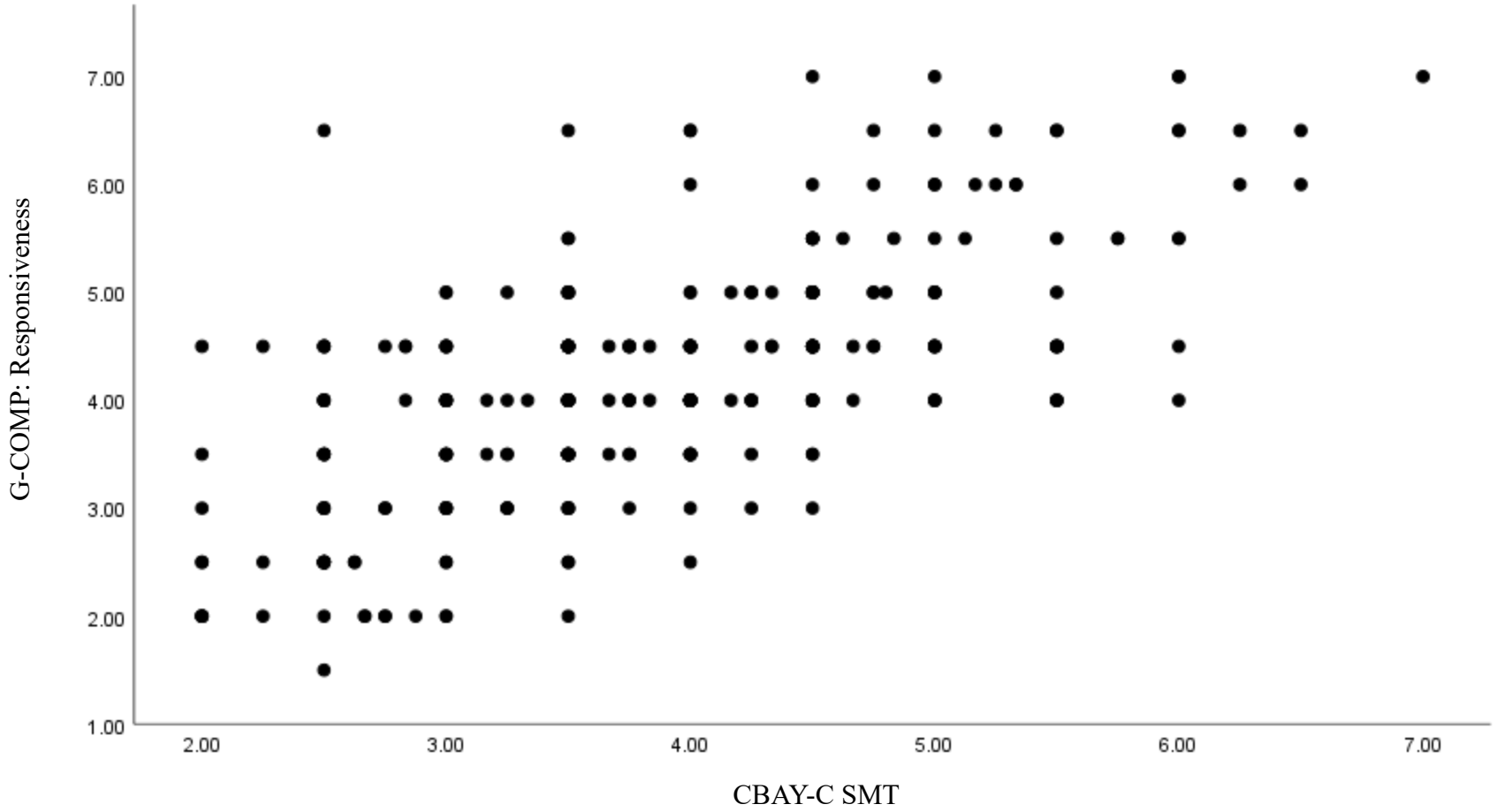

Figure 10. Correlation between CBAY-C SMT and G-COMP: Responsiveness subscale scores in the SMT condition.

Appendix

Cognitive-Behavioral Treatment for Anxiety in Youth Competence Scale (CBAY-C)

Instructions: Using the grid provided below, please indicate COMPETENCE of any item observed for each ten-minute time segment. Specify all target problems that apply by circling the appropriate Problem Items ("A" = Anxiety, "D" = Depression, "C" = Conduct). Use a "+" to indicate an above average rating, "X" to indicate an average rating, and "-" to indicate a below average rating. After watching the ENTIRE recording, use the 1-7 scale to assign a Competence rating (Comp) for all items that are present in at least ONE (1) time period. Check the box below if you think the primary target problem is Anxiety.

☐ Primary problem is Anxiety, ☐ Primary problem is Depression, ☐ Primary problem is Conduct

	0 Not Present	1 Very Poor	2 Poor	3 Acceptable	4 Adequate	5 Good	6 Very Good	7 Excellent		
	Item								**Comp**	
Standard	1. General Technical Comp									1.
	2. Model Structure/Phase									2.
	3. HW Review									3.
	4. HW Assigned									4.
	Problem Item(s): A D C									
Model	5. Psychoeducation									5.
	6. Emotion Ed									6.
	7. Fear Ladder									7.
	8. Relaxation									8.
	9. Cognitive									9.
	10. Problem Solving									10.
	11. Self-Reward									11.
	12. Coping Plan									12.
	13. Exposure Preparation									13.
	14. Exposure									14.
	15. Debriefing an Exposure									15.
	16. Maintenance									16.
	17. Activity Selection									17.
	18. Talents and Skills									18.
	19. Positive Self									19.
	20. One-on-One Time									20.
	21. Praise									21.
	22. Active Ignoring									22.
	23. Effective Instructions									23.
	24. Rewards									24.
	25. Time Out									25.
	26. Daily Report Card									26.
Delivery	27. Didactic Teaching									27.
	28. Collaborative Teaching									28.
	29. Modeling									29.
	30. Rehearsal									30.
	31. Coaching									31.
	32. Self-Disclosure									32.
	33. Play/Creative									33.

Global Therapist Competence Scale for Youth Psychosocial Treatment (G-COMP)

Instructions: Using the grid provided below, please indicate whether each specific item occurs during each ten-minute time segment. If an item occurs during a time segment place a "+" to indicate above average rating, "X" to indicate average rating, or a "–" to indicate below average rating in the space provided in the grid corresponding to the correct item. After watching the ENTIRE recording, use the 1-7 scale to assign a Competence (Comp) rating for all items that are present in at least ONE time period.

1 Very Poor	2 Poor	3 Acceptable	4 Adequate	5 Good	6 Very Good	7 Excellent

MICROANALYTIC ITEMS									Comp
1. Alliance Building									
• **Conveys understanding/validates client's experience:** ability to identify, understand, and accurately communicate an understanding of the client's behavior, affect, and feelings. • **Demonstrates positive regard:** ability to display interpersonal skills (warmth, empathy, congruence, genuineness) that promote/maintain the affective component of the client-therapist relationship. • **Elicits client's perspective/experience:** ability to elicit and remain focused upon the client's subjective perspective and experience. • **Fosters collaboration:** ability to make therapy, including setting tasks and goals, a collaboration that is equally shared between therapist and client.									
2. Positive Expectancies									
• **Facilitates treatment expectancies:** ability to establish the credibility of the treatment model and promote client expectations that the treatment model is the right fit for the client (instillation of hope). • **Facilitates therapist credibility:** ability to establish for the client that he/she has the necessary knowledge and expertise to successfully treat the client's presenting problem. • **Facilitates client self-efficacy:** ability to instill the belief within the client that he/she is able to perform therapeutic activities and achieve treatment goals.									
3. Focusing Treatment									
• **Structure and Pace:** ability to provide clear structure and organization to a session with minimal departures from session focus/theme. • **Continuity of treatment:** ability to connect content across sessions to create a sense that treatment is a coherent whole and NOT a series of single sessions. • **Focuses on key themes in session:** ability to maintain focus on a small set of themes or ideas during a therapy session.									
4. Instigating Change									
• **Uses change strategies effectively:** ability to deliver effective therapeutic interventions to address the cognitive, behavioral, and/or affective elements of the client's presenting problem. • **Facilitates active client participation:** ability to create a session that is characterized by mutual client-therapist activity, interest, and effort directed at therapeutic activities.									
5. Responsiveness									
• **Fosters client's motivation:** ability to deal effectively with client resistance, low motivation, or disagreements with treatment goals. • **Flexibility:** ability to tailor the content and structure of treatment to each client.									

Vita

Jennifer Cecilione was born on January 24, 1993 in Long Branch, New Jersey. She graduated summa cum laude from the University of Richmond in May 2015 with a Bachelor of Arts in Psychology. She then worked for three years as a research coordinator at the Virginia Institute for Psychiatric and Behavioral Genetics before entering the Clinical Psychology doctoral program at Virginia Commonwealth University in Richmond, Virginia.