2017

VALIDITY OF CLINICIANS' SELF-REPORTED TREATMENT TARGETS ON THE MONTHLY TREATMENT PROGRESS SUMMARY

Allison K. Powell
University of Montana, Missoula

VALIDITY OF CLINICIANS' SELF-REPORTED TREATMENT TARGETS ON THE MONTHLY TREATMENT PROGRESS SUMMARY

By

ALLISON KAORU POWELL

Bachelor of Arts, University of Hawaii, Honolulu, HI, 2011

Thesis

presented in partial fulfillment of the requirements
for the degree of

Master of Arts
in Clinical Psychology

The University of Montana
Missoula, MT

July 2017

Approved by:

Scott Whittenburg, Dean of The Graduate School
Graduate School

Christine Fiore, PhD, Co-chair
Psychology Department

Cameo Stanick, PhD, Co-chair
Hathaway-Sycamores Child and Family Services

Duncan Campbell, PhD
Psychology Department

Rick van den Pol, PhD
Department of Curriculum and Instruction

Powell, Allison, M.A., Summer 2017 Clinical Psychology

Validity of clinicians' self-reported treatment targets on the monthly treatment progress summary

Co-Chairperson: Christine Fiore, PhD

Co-Chairperson: Cameo Stanick, PhD

Several decades of research have been spent identifying and testing EBTs, but there is currently very little research that examines the therapeutic practices within usual care. The lack of understanding in this area has been implicated as a factor which hinders the successful implementation of evidence-based therapies (EBTs) into usual mental health care settings. The Monthly Treatment Progress Summary (MTPS) is a measure developed to enable monthly tracking of intervention strategies and content within a statewide system of children's mental health care. Although a growing body of research exists examining the reliability and validity of the treatment practice and progress sections of the MTPS, less research has been conducted on the treatment target section. Specifically, no information exists on the validity of the treatment target section, which is a significant limitation of the measure and has implications for gaining a full picture of usual care treatment. The current study demonstrated some support for the use of the MTPS as an indirect measure of specific content on which clinicians focus in therapy. Trained coders were able to reliably identify specific treatment targets focused on in treatment sessions for 12 of 13 targets that occurred at a high enough frequency to be analyzed. Overall coder-clinician agreement was low, with four of the 12 targets achieving acceptable levels of reliability (ICC $\geq$ .60). These results suggest there may be a difference between clinician intent and the observable content a clinician engages in. These results may also indicate differential levels of familiarity or understanding between coders and clinicians of what constitutes 'focus' in session, as the present study demonstrates that several treatment targets on the MTPS can be reliably coded. Future research should incorporate larger samples of more diverse clients that will include content related to treatment targets that were removed from the present analysis. Lastly, the discrepancies noted between coders and clinicians indicate the need for future research to elucidate clinician intent, clinicians' accuracy in reporting their session content, and the relationship between clinician intent and observable session content.

Introduction

Usual Care Studies & Variability in Usual Care

Within the past two decades, there has been increased attention placed upon the implementation of evidence-based treatments (EBTs) into the therapeutic practices of usual care (UC) (Fixsen, Naoom, Blase, & Friedman, 2005; Weisz, Jensen-Doss, & Hawley, 2006). However, the dissemination of EBTs into UC has been a slow advancement beset with many difficulties (Balas & Boren, 2000; Kelley, Vides de Andrade, Ana Regina, Sheffer, & Bickman, 2010). In order to provide the best care available for consumers, the dissemination and implementation of effective mental health services should be a priority for service organizations (Aarons, Hurlburt, & Horwitz, 2011; Fixsen et al., 2005). Understanding the challenges that can occur during implementation will help guide improvements in the implementation process.

One of the largest challenges for dissemination and implementation science is that, despite knowledge of EBTs, many clinicians in UC do not utilize them (Bickman, 2008b; Borntrager, Chorpita, Higa-McMillan, & Weisz, 2009; Hembree & Cahill, 2007). Much of this reluctance has been attributed to the erroneous beliefs that many clinicians have about EBTs (Hembree & Cahill, 2007). For example, concerns have been raised over the efficacy of EBTs with the more severe and complex clinical cases seen in UC populations, based on the assertion that participants in randomized trials, from which EBTs are developed, are straightforward treatment cases with single diagnoses and no significant comorbidity (Weisz et al., 2006). Additionally, EBTs have been criticized as inflexible and therefore of limited use in a diverse population with varied therapeutic needs.

Several decades of research have been spent identifying and testing EBTs, but there is currently very little research that examines the therapeutic practices within UC. Many researchers have referred to this paucity of knowledge as a "black box" of which very little is

known or understood (Bickman, 2000; Garland, Bickman, & Chorpita, 2010; Kelley et al., 2010). Multiple obstacles exist that impede research in UC settings. Psychotherapeutic interactions are complex and contain inherent difficulties with regards to obtaining reliable and accurate measurements of therapeutic processes and outcomes. However, such information is imperative to characterizing comparisons with EBTs, as well as in identifying potential barriers to successful implementation of EBTs (Garland, Hurlburt, Brookman-Frazee, Taylor, & Accurso, 2010). Other researchers have attributed the difficulties in characterizing UC and the application of EBTs in UC to the discrepancy between the focused nature of EBTs and the breadth of factors that affect treatment involvement in UC (Aarons et al., 2011; Southam-Gerow, Chorpita, Miller, & Gleacher, 2008). One aspect in particular that has been frequently discussed is differences between the settings of and populations seen in UC and in research (Baker-Ericzén, Hurlburt, Brookman-Frazee, Jenkins, & Hough, 2010; Garland, Bickman, et al., 2010; Southam-Gerow et al., 2008). Research evidence from which EBTs are derived, is typically gathered in research clinics; whereas, UC is typically performed in public community settings, such as schools, in homes, or community mental health centers. Research clinics are often focused on the treatment of one specific problem area, and utilize referral systems that are targeted towards clients whose psychopathologies are dominantly characterized by their problem area of focus (e.g. advertisements and professional referrals). Conversely, community clinics provide services for a broader range of psychopathologies and rely on more varied sources of client referrals, including government agencies, insurance, and schools. Research investigating the differences within the client population that each setting serves has shown that clients in UC have higher rates of comorbidity and are more ethnically diverse (Garland, Bickman, et al., 2010). Southam-Gerow et al. (2008) observed significantly higher levels of psychosocial

stressors and lower family income for clients in UC settings than in research settings. However, despite the documentation of these differences, the extent to which treatment outcomes are affected is not known.

In order to better understand the sources of variability within UC, several areas have been investigated. Lutz, Leon, Martinovich, Lyons, and Stiles (2007) conducted a multi-level analysis within a naturalistic data set of 1,198 patients and 60 clinicians to assess for the amount of variance in across-session change in symptom intensity explained by therapist differences. The analyses included three levels of nested factors (sessions within patients, patients within therapists, and therapists) and controlled for differences among patients in therapist-rated and patient-rated symptom intensity, therapeutic expectations, symptom chronicity, and previous treatment experience. The authors found that 17% of the variance in patients' rate of improvement in across-session symptom intensity was explained by individual differences between therapists. This finding was substantially higher than values found in similar studies, which reported values of about 8% (Kim, Wampold, & Bolt, 2006; Wampold & Brown, 2005). This difference suggests that the real-world sample used by Lutz et al. (2007) contained a wider range of therapist skill, knowledge, and personality than controlled clinical trials, as studies that standardized treatment using manualized therapies and more experienced therapists showed the smallest therapist effects (Elkin, Falconnier, Martinovich, & Mahoney, 2006). The differences found in the amount of variance due to therapist individual differences by Lutz et al. (2007) and studies that examined clinical trials underscores the need for accurate methods of characterizing therapist behaviors in UC.

Efforts to describe UC have focused on identifying and evaluating therapist practices; however, these studies have helped explain only a portion of the characteristics of UC (Bickman,

2008a). Other aspects of psychotherapeutic services in UC must be researched to identify other factors that contribute to the characterization of UC. For example, much of clinical research is driven by diagnoses; however, within UC, a diagnosis might not accurately capture the reasons clients seek psychotherapy, or the areas of improvement that clients desire as an outcome of their treatment experience (Daleiden, Lee, & Tolman, 2004).

One of the most frequently used diagnoses in UC settings is the broad *not otherwise specified* (NOS) (Widiger & Samuel, 2005). Clinicians generally use NOS when they have determined that a disorder is present, but the client does not meet the diagnostic criteria. The frequent use of NOS suggests that clinicians view diagnoses as being of limited clinical utility, and are not necessarily indicative of particular targets of treatment (Clark, Watson, & Reynolds, 1995; Verheul & Widiger, 2004). This indicates a need for an alternate or complementary system for describing clients' presentations that is capable of capturing the diverse range of symptoms seen in UC settings.

One method for accomplishing this is through the establishment of treatment targets. Nezu and Nezu (1993) describe treatment targets as the outcome of a process in which "the therapist translates a client's complaints of distress into a meaningful set of target problems and treatment goals" (p. 254). Though correlated with diagnoses, treatment targets are generally determined by the client and represent the improvements the client desires as a result of their treatment involvement (Daleiden et al., 2004; Nezu & Nezu, 1993). Nezu and Nezu (1993) described two types of goals: ultimate outcomes and instrumental outcomes. Ultimate outcomes address the overall reason for seeking treatment, whereas the instrumental outcomes represent goals that are targeted in order to achieve the ultimate outcome. For any ultimate outcome, there is a wide range of instrumental outcomes that therapists might target during treatment.

Similarly, although one can conceptualize individual treatment targets as symptoms that cumulatively represent categorical diagnoses, treatment targets allow for additional areas of clinical focus that are not disorder specific (e.g., family functioning, peer relationships)(Child and Adolescent Mental Health Division, 2003). This perspective is particularly beneficial to UC settings, in which rates of comorbidity among youth have been observed at 50%-75% (Higa-McMillan, Powell, Daleiden, & Mueller, 2011; Weersing, Iyengar, Kolko, Birmaher, & Brent, 2006). Previous research has suggested that individuals with comorbid diagnoses respond differently to treatment than individuals with a single diagnosis (Hurlburt, Garland, Nguyen, & Brookman-Frazee, 2010; Kessler, Chiu, Demler, & Walters, 2005). However, current categorically-based diagnostic measurements of treatment make it difficult to evaluate the potential differences in treatment response among those with comorbid diagnoses (Helzer, Kraemer, & Krueger, 2006). By focusing on the specific problem areas that clients experience, therapists can create stronger ties between the practices and outcomes.

The use of treatment targets to conceptualize clients' clinical problems also presents an advantage in the ability to measure clients' progress in therapy (Love, Tolman, Mueller, & Powell, 2014). Treatment targets can be examined both individually and aggregated together, which allows for new ways to assess and analyze outcomes in UC studies. Measuring treatment outcomes at the treatment target level can provide detailed information about a client's response to treatment, in relation to areas where a client may be making gains more quickly or be experiencing a slower rate of improvement. Additionally, in order to address the limitations associated with the use of categorical diagnoses in outcome studies, symptoms related to disorders can be conceptualized as individual treatment targets (Love et al., 2014).

Importantly, less attention has been afforded to treatment targets in the literature, compared to research on treatment practices. In general, research examining UC practices has demonstrated that UC therapy is less structured, less behavioral, and more eclectic than most manualized EBTs (Weersing, Weisz, & Donenberg, 2002). In one study of UC treatment for youth with disruptive behavior problems, Garland, Brookman-Frazee, et al. (2010) observed 218 children ages 4-13 who were receiving services in six publicly-funded clinics for 16 months. The authors observed a large amount of variability within the amount and the type of treatment provided. Children attended an average of 22 sessions (range 0-63 sessions). A sample of 1215 videotaped sessions were randomly selected and coded for 27 practice elements. The authors found that therapists engaged in many different practice elements, that is discrete therapeutic techniques, both within and across therapy sessions, and targeting both the children and their caregivers. However, the intensity of the delivered practice elements was low (2.3 on a scale of 1-6). What is guiding UC clinicians' decisions to engage in a breadth of techniques, even at the expense of depth, is unclear however. Lack of training in specific techniques, responding to crises, addressing comorbidity may all be factors, but in general an important next step in understanding UC and successfully implementing EBTs involves researching the drivers behind clinicians' practice behaviors.

Current Methods of Research into Treatment Practices

Given what little is known about the nature of UC therapy, the need for a system of measurement in UC is clear (Bickman, 2008a; Garland, Hurlburt, et al., 2010). In addition, with the increased utilization of managed care, medical providers are increasingly required to document treatment processes and outcomes (Nakamura, Daleiden, & Mueller, 2007). Further, increased national attention on the dissemination and implementation of EBTs has required greater focus on accountability within UC.

Currently, two methods exist for examining treatment practices. The first is through direct observation, in which clinicians are supervised during treatment with clients. This can be achieved through video or audio recordings of therapy sessions, or through in vivo observation. The second is through indirect measurements, including clinician self-report measures or reviews of treatment session notes. Measures such as the Therapy Process Observational Coding System-Strategies scale (McLeod & Weisz, 2010), the Therapy Procedures Checklist (Weersing et al., 2002), and the Monthly Treatment Progress Summary (Child and Adolescent Mental Health Division, 2003) have been developed to assess therapy according to these two methods. In particular, these measures have been used to determine fidelity in the use of EBTs within UC settings and have provided valuable information regarding UC therapeutic practices (Borntrager, Chorpita, Orimoto, Love, & Mueller, 2013; Garland, Brookman-Frazee, et al., 2010). However, these measures have limitations regarding their feasibility of use and the accuracy with which they characterize UC. Research investigating the effectiveness of EBTs in UC settings is hindered by our current inability to sufficiently document or measure, and therefore understand, the practices within UC (Kelley et al., 2010). In order to bridge the research-practice gap between EBTs and UC, we must first understand the current level of functioning of 'treatment as usual,' and identify what areas are in need of improvement (Garland, Bickman, et al., 2010).

The Therapy Process Observational Coding System-Strategies scale (TPOCS-S) is one example of a measurement using direct observation (McLeod & Weisz, 2010). The TPOCS-S is a coding system that was developed to provide an objective description of UC. The measure is comprised of 31 items that cover a wide range of therapeutic practices from multiple treatment areas. The authors derived four subscales from a self-report measure of therapeutic practices, the Therapy Procedures Checklist (TPC) (Zoffness, Garland, Brookman-Frazee, & Roesch, 2009) to

represent relevant treatment domains in child therapy: Cognitive, Behavioral, Psychodynamic, and Family. The authors developed a fifth subscale, called the Client-Centered subscale, based on the frequent endorsement of client-centered interventions by child therapists. Items from the TPC and the Therapist Behavior Rating Scale (TBRS) that represented observable therapist behaviors were identified and combined into the current TPOCS-S scale. The authors also added 6 new items that represented often used interventions that were not associated with a specific treatment domain (e.g., play-therapy, homework, treatment goals).

In order to utilize the TPOCS-S, therapy sessions are videotaped and scored in 5 minute increments. The scoring criteria for the TPOCS-S uses an 'extensiveness' rating as well as a measurement of 'frequency.' The extensiveness rating scale indicates the thoroughness with which a therapist engages in a therapeutic practice on a 7-point Likert scale ranging from 1 (*not at all*) to 7 (*extensively*). Frequency refers to how often an intervention was used within one session.

The reliability and validity of the TPOCS-S was assessed using a sample of 43 children across six community mental health clinics who were receiving UC treatment for internalizing disorders. Videotaped therapy sessions for participating children were coded for frequency and extensiveness of interventions. The TPOCS-S showed acceptable internal consistency and validity, with intraclass correlations (ICCs) ranging from 0.51 (fair agreement between raters) to 0.94 (excellent agreement between raters) (McLeod & Weisz, 2010).

The TPOCS-S is a promising measurement for characterizing therapeutic practices within UC. However, it is a time and labor-intensive measurement that may not be feasibly implemented in healthcare settings which do not possess the resources to ensure their completion. The process of videotaping and incrementally coding each session is likely to be too

cumbersome, time-consuming, expensive, and intrusive to be readily and widely adopted among community-based healthcare providers. Additionally, observational methods do not account for or measure therapist intent (Weersing et al., 2002). Weersing et al. (2002) describe therapeutic techniques as involving both an observable action and an underlying purpose for the action. In measures that rely on direct observation, information regarding the purpose of actions is difficult to obtain. Information from the clinicians themselves about their therapeutic techniques must be obtained as well.

In community-based healthcare settings, indirect measurements of clinical practice have several advantages. The use of an indirect measure, such as clinician self-reports, lessens the burdens associated with resources such as time, finances, and staffing (Borntrager, Chorpita, Orimoto, et al., 2013). The Therapy Procedures Checklist (TPC; Weersing et al., 2002) is an example of an indirect measurement of UC practices. The TPC is a 62-item clinician self-report measure of therapeutic techniques that contains three subscales associated with the most common theoretical orientations used in therapy with children: Cognitive, Behavioral, and Psychodynamic. Items were created to represent central and specific interventions for each of the orientations. The authors collected techniques through reviews of child psychotherapy literature and refined the items using reviews by the authors as well as by other child psychotherapy researchers and therapists in community mental health centers. Techniques that were observed within multiple domains were removed. The TPC has demonstrated good internal consistency, test-retest reliability, and sensitivity to between-client changes in technique use in a sample of 16 therapists and 108 child clients from community mental health clinics in Los Angeles; however, the TPC is limited as a measure of practices and does not allow for measurement of other features of therapeutic intervention (e.g., treatment targets, progress, etc.)

Monthly Treatment Progress Summary

Another indirect measure for characterizing clinician's behavior that has a growing body of evidence to support its psychometric and clinical properties is the Monthly Treatment Progress Summary (MTPS; Child and Adolescent Mental Health Division, 2003), which was developed to enable monthly tracking of intervention strategies and content within a statewide system of children's mental health care (Chorpita & Donkervoet, 2005). The State of Hawaii Child and Adolescent Mental Health Division (CAMHD) requires providers to report monthly summary information for all youth registered in the system using an empirically-derived checklist of treatment objectives and practices to report treatment practices, treatment targets, and progress towards treatment goals (Child and Adolescent Mental Health Division, 2003; Daleiden et al., 2004).

On the MTPS, ''treatment targets'' refer to therapeutic objectives that incorporate both diagnostic symptoms and non-diagnostic treatment goals. During the treatment planning phase, treatment targets are collaboratively identified between the service provider and the family receiving services. The identified targets are related to, but distinct from diagnoses, which are not conceptualized as the focal point of treatment engagement (Daleiden & Chorpita, 2005). Diagnoses serve as a factor representing common treatment goals and guiding treatment intervention selection. Treatment targets represent an effort to characterize the multi-faceted needs of youth engaged in UC.

"Treatment practices" refer to the intervention strategies or practice elements used by therapists. The items in the practice element section of the MTPS represent both practices found in evidence-based literature and practices common in UC that were not necessarily part of the evidence-based literature.

The items in the treatment target and treatment practice sections of the MTPS were derived from a comprehensive review of evidence-based literature for youth mental health treatment (Daleiden & Chorpita, 2005). The treatment target section of the MTPS contains a list of 48 predefined targets, plus two open-response targets, from which providers endorse up to 10 items that were the focus of treatment during the reporting month. The practice elements section is comprised of 55 predefined items and three open-response fields. There is no limit on the number of practice elements a provider may endorse.

Previous studies of the MTPS have demonstrated support for the reliability and validity of the treatment targets, practice elements, and progress ratings. Daleiden et al. (2004) found therapist endorsement of treatment targets to be significantly related to the diagnoses of youths at intake, which provides preliminary support for the convergent and discriminant validity of endorsed treatment targets, though other forms of validity have yet to be established. Nakamura et al. (2007) found that therapist-reported change in progress ratings on the MTPS were significantly related to client change, defined as decreased scores on measures of functional impairment, which provides support for the validity of progress ratings as a measure of client change in functional status over time. Borntrager, Chorpita, Orimoto and colleagues (2013) coded audio-recorded therapy sessions for practice elements and compared the results to the therapist-endorsed practice elements reported on the MTPS. The authors found that, for a subset of reliably coded practice elements, the practice element section of the MTPS provided a valid measure of clinician behavior within treatment sessions. Although a growing body of research exists examining the reliability and validity of the treatment practice and treatment progress sections of the MTPS, less research has been conducted on the treatment target section. Specifically, no information exists on the validity, besides convergent/divergent validity, of the

treatment target section, which is a significant limitation of the measure and has implications for gaining a full picture of the 'black box' of UC treatment.

Hypothesis

The purpose of the current study is to assess the reliability and the validity of the treatment target section of the MTPS. Reliability was assessed via examining intraclass correlations between a team of coders to determine whether treatment targets could be reliably coded. Validity was assessed via the comparisons of coder-rated treatment targets to the self-reported treatment targets of clinicians, also via intraclass correlation analyses. These analyses were used to evaluate whether clinicians' self-reported treatment targets are consistent with their treatment objectives in session as judged by trained coders reviewing audiotapes of therapy sessions. A strong relationship between coder and clinician ratings would indicate that the treatment target section of the MTPS is capturing a valid representation of what clinicians focus on in treatment.

Methods

Data for the current study were collected for the month of April 2008, during a randomized clinical effectiveness trial examining children's mental health treatments (Child Systems and Treatment Enhancement Projects [Child STEPs] (Chorpita et al., 2013; Weisz et al., 2012). Full details of the Child STEPs recruitment process are described elsewhere (Chorpita et al., 2013; Weisz et al., 2012); although, participant and methodology information relevant to the current study are described below.

Participants

Clinician Participants

Audio recordings of therapy sessions for the month of April 2008 and the corresponding MTPS forms were collected from 16 therapist participants at the Hawaii site of the Child STEPs trial. One audio recording, which was the only session submitted by the clinician, was inaudible, thus the resulting data was taken from a total of 15 clinicians. Clinicians ranged in age from 26 to 59 years, with a mean age of 38.3 years. Their years in practice ranged from 2 to 21 years, with a mean of 7.2 years. Sixty percent of the clinicians reported their ethnicity as White. Forty percent of the clinicians reported their ethnicity as Asian American. The majority of clinicians held a Master's degree as their highest obtained degree (86.7%), whereas 13.3% held a doctorate degree. Table 1 shows all demographic characteristics for clinicians.

Clinicians were recruited for the Child STEPs trial across school- and clinic-based mental health settings and private practices. For the Child STEPs trial, clinicians were randomized into either one of two EBT groups or a UC group. Clinicians in the EBT groups received training. In the UC group, clinicians were instructed to provide treatment as they would outside of the study. For the duration of the Child STEPs trial, clinicians audio recorded all of their treatment sessions, a month's worth of which will be used for the current study.

Child Participants

Data was collected from 20 children who were in the active treatment phase of the Child STEPs trial during the 1 month period (April 2008). As described above, one child's session recording was inaudible, resulting in a total of 19 children. The children ranged in age from 7 to 13 years (M=9.55, SD=4.24). Their reported ethnicities were 55% mixed race, 35% White, 5% African American, and 5% Asian American. A total of 47 sessions were collected, inclusive of all 3 treatment conditions, with youth ranging from one to five session during the month (M=2.7, SD=2.83). Sessions were conducted in a variety of settings, including schools and clinics. Table

2 shows all demographic characteristics for youth and Table 3 shows the number of sessions recorded per client.

Child participants were referred by clinicians participating in the Child STEPs trial. Appropriateness for the study for referred youth was completed through a telephone screening process. After consent was obtained, youth and their caregiver(s) were administered a standardized diagnostic assessment. Youth with a primary diagnosis of anxiety, depression, or disruptive behavior were included in the study. Youth with mental retardation, current suicidal ideation, or psychosis were excluded from the study. Final study inclusion was determined by the project team.

Coders

The coding team consisted of two doctoral students in clinical psychology (including first author), one postdoctoral fellow in clinical psychology, and one clinical psychology faculty member serving as the expert coder. Two of the coders are Caucasian, including the expert coder, and two are Asian American. They range in age from 27 to 37. All but one coder had previous coding experience, and coders had between 1 to 8 years of graduate training in clinical psychology.

Measure: Monthly Treatment Progress Summary

Monthly Treatment Progress Summary (MTPS; Child and Adolescent Mental Health Division, 2003). As described above, the MTPS is a clinician self-report measure which tracks treatment goals, progress ratings, and treatment practices on a monthly basis. Forty-eight treatment targets and two-open response field constitute the treatment target section. The MTPS form and codebook are available via the CAMHD website (Child and Adolescent Mental Health

Division, 2008a; Child and Adolescent Mental Health Division, 2008b). The MTPS form is shown in Appendix A.

Preliminary support for the reliability and validity of the treatment target section of the MTPS has been demonstrated in previous studies. Target endorsement was found to be significantly related to the diagnoses of youths at intake, which provides preliminary support for the convergent and discriminant validity of endorsed treatment targets (Daleiden et al., 2004). Nakamura, Daleiden, and Mueller (2007) found moderate overall month-to-month stability (k=0.66) and three-month stability (k=0.52). These results support the reliability of the MTPS treatment target section throughout treatment.

Procedure

Coding Training

All coders had previous experience using the MTPS in either or both a clinical or research context. Coders underwent extensive training in the coding system and were allowed to refer to the MTPS codebook when listening to sessions. During the training period, coders completed practice coding sheets on therapy sessions that were collected outside of the target month. Coders underwent training until excellent interrater reliability was reached (an intraclass correlation [ICC] of 0.90 or higher; Cicchetti & Sparrow, 1981).

Coding Procedure

A two-level coding system for providing *categorical* and *focus* ratings was developed for identifying treatment targets. First, treatment targets were organized under the factors outlined in Love, Okado, Orimoto and Mueller (2016). Love et al. (2016) identified a five-factor model which represented patterns of clinician-endorsed treatment targets in an archival data set of MTPS ratings. The subset of treatment targets that demonstrated adequate loadings in the factor

analysis were chosen to represent treatment targets in the current coding system. 'Focus' ratings reflect the extensiveness/amount of focus a factor (and the subsumed targets) was given. A target would be coded as a "0" if it was in no way targeted or associated with session content. A "1" would be coded if a clinician made a brief statement or action that implied a connection with a target (ex. "You were very brave" as the only contribution from clinician associated with client discussing something related to the target *Phobias/Fears*). A "2" would be coded if, at minimum, a clinician engaged in an explicit discussion or activity that clearly addressed a treatment target by target name, a synonym for the target name (ex. "sadness" as a synonym for the target *Depressed Mood*), or by the exact definition of a target (as defined in the MTPS Coding Manual, see Appendix B). These ratings were completed at the treatment target level. Also, if greater than or equal to 50% of the treatment targets under a factor received a "1" or higher, a "2" would be coded for that factor as whole. These focus ratings at both the individual treatment target and factor level allowed for greater depth in understanding clinicians' intentions in their treatment sessions.

Categorical ratings were also coded, reflecting the presence/absence of a treatment target in the diagnostic profile of a participant. Following the initial, 'blind' focus coding, coders were provided the diagnostic profile information for each youth participant, to which participating clinicians were also privy. A "0" was coded if no treatment target was identified. A "1" was coded if a treatment target was identified within the diagnostic profile. Only a subset of treatment targets were clearly related to specific diagnoses (*Oppositional/Non-Compliant Behavior*; Oppositional Defiant Disorder). Tables 4 and 5 show the full treatment targets listed and possible, codeable diagnoses.

Following the training period, two coders from the coding team were randomly assigned to each child participant's recorded therapy sessions. One coding sheet was completed, per coder, for each session. Microanalytic coding sheets were organized into 5-minute increments to allow for easy identification of code discrepancies (McLeod & Weisz, 2010). Discrepancies were reconciled by the third, expert coder to create a third coding sheet per session. Coding sheets were condensed per participant, in the event that there was more than one session in the target month. The highest focus rating given per treatment target, per participant, was used to calculate ICCs among coders to establish treatment targets that could be reliably coded at ICCs ≥ .60, which previous research has identified as a moderate degree of reliability (Borntrager, Chorpita, Orimoto, et al., 2013; Cicchetti & Sparrow, 1981; McLeod &Weisz, 2010).

For comparisons between the coding team and clinicians' ratings, the 'final' focus coding ratings for each treatment target that received a "1" or higher were recoded as a "1" and everything else remained a "0". These ratings were compared to clinicians' endorsements of treatment targets for final ICCs. The recoding adjustment of 2 → 1 and 1 → 1 ratings was necessary given that clinicians, per the MTPS instructions, completed endorsed treatment targets dichotomously (targeted/not targeted). The initial coding of 0-2 created a less conservative estimate of clinicians' intentions regarding their treatment focus. In a study examining clinician-reported therapeutic practices, Borntrager, Chorpita, Orimoto, and colleagues (2013) found that when a more stringent coding system was used (0/1; absent/present) coder-clinician agreement was low (33.33%); however, when a coding system with more allowable variance was utilized (0 = no explicit or implicit use; 3 = at minimum, explicit discussion or behavior), coder-clinician agreement was substantially higher (100%). In addition, one clinician-rated MTPS form

contained 13 endorsements for treatment targets. As the instructions for this form are to mark up to 10 targets, the first 10 targets were used in analyses and the last three were excluded.

Lastly, factor ratings were explored between the coding team and clinicians. In order to make these comparisons, only treatment targets that loaded onto one of the five factors were analyzed from the clinicians' MTPSs. If greater than or equal to 50% of the treatment targets under a factor were endorsed, a "1" was assigned for that factor. These factor ratings were compared to the coder-endorsed focus ratings.

Analyses

Following the training period, intra-class correlations (ICCs) were calculated among the coders for each factor/treatment target (Shrout & Fleiss, 1979). Those treatment targets that resulted in ICCs ≥ .60 were used for comparison to clinicians' MTPSs. To examine coder-clinician agreement and thus establish validity, the clinician-reported MTPSs were then compared to the final scores obtained from coders (converted to dichotomous ratings as described above).

Results

The rate of endorsement between coders and clinicians was compared using means and standard deviations calculated across factor-loaded treatment targets. On average, clinicians (M=3.36, SD=1.84) endorsed more targets than the coding team (M=2.89, SD=1.65). Tables 6 and 7 show the frequency counts of total treatment targets per client and the frequency counts of available treatment targets between coders and clinicians.

Of the 32 treatment targets available for coding based on the factor analysis described previously (Love et al., 2016), 11 treatment targets were removed from analyses because they were not endorsed or were endorsed with such low frequency that they could not be reliably

coded. These included any targets that were endorsed via the write-in line. The model ICC (2,3) was used to assess intraclass correlations, which indicated that coders were randomly assigned to audiotapes, and that reliability was measured by taking the average rating of the three coders. For the remaining 13 treatment targets within the factor structure described above, the coding team established acceptable interrater reliability (ICCs of 0.60 or higher) for 92% of available treatment targets (Table 8). The only target that did not achieve acceptable interrater reliability was hyperactivity (ICC=0.589), though it was approaching acceptability. These 12 targets were retained for comparison between coders and clinicians.

Importantly, the expert coder examined coding discrepancies to assess for drift and to resolve discrepant codes. The expert coder did not find evidence that indicated a pattern of drift for any coders, such as one coder systematically endorsing targets more conservatively or leniently than other coders. The expert coder assessed discrepancies and altered codes in <1% of available codes.

In order to evaluate whether clinicians' self-reported treatment targets were consistent with their treatment objectives in session, as judged by trained coders reviewing audiotapes of therapy sessions, and thus establish validity, the researcher reduced the final ratings obtained from coders to dichotomous ratings to facilitate coder-clinician comparison. The researcher assessed intraclass correlations using the model ICC (3,2), which indicated that final coding ratings were compared with each clinician. Of the 12 targets that were compared to clinicians' ratings, four targets (33.33%) had an ICC of 0.60 or higher. These targets were *Aggression*, *Anger*, *Oppositionality*, and *Phobias/Fears* (Table 8). In addition to the overall low rate of consistency between coders and clinicians, it is important to note that one target resulted in a negative ICC value (attention problems = -0.46). This value indicates that the ratings provided

by coders vs. clinicians were systematic in disagreement (Hallgreen, 2012). For this target, clinicians coded for the presence of the target (i.e. that attention problems were a focus of session) for sessions where coders did not endorse the target, and for sessions in which coders endorsed attention problems, clinicians did not.

Lastly, the study examined categorical ratings, which were acquired through the diagnostic profile of each participant as described above, and compared the ratings to the clinicians' self-reported targets. The researcher assessed these comparisons using the model ICC (3,2), which represented the categorical ratings compared to clinicians' ratings. Of the original 15 targets and associated diagnoses, 10 were endorsed at least once through participants' diagnostic profiles. From those 10, only two targets (20%) resulted in an ICC of 0.60 or higher, which were *Eating, Feeding Problems* and *Anxiety* (Table 9).

Discussion

The current study provides reliability and validity information about the treatment target section of the MTPS, a clinician-report measure used within a statewide healthcare system to track mental health service provision. A number of treatment targets were not endorsed or were endorsed at a frequency too low to analyze; however, for targets that could be reliably coded, adequate interrater reliability was achieved for 12 of the 13 targets. For these 12 targets, coder-clinician agreement was low, but reached adequate reliability on four targets (33.33%). Although these results indicate overlap between the treatment targets clinicians endorsed and the targets discerned by trained coders, the discrepancies between the overall endorsements highlight the need for further attention to the reporting practices of clinicians identifying the targets on which prior sessions were focused.

Prior research on usual care (UC) practices indicate that clinicians tend to overreport their therapeutic practices rather than underreport (Garland et al., 2010 & Hurlburt et al., 2010). The results of this study parallel this literature, in that, on average, clinicians reported slightly more targets which were a focus over the course of one month of therapeutic services than coders. Additionally, Garland and colleagues (2010) observed that clinicians tended to address different therapeutic content with children and with parents. It is possible, though speculative, that clinicians reported targets that represented their work with parents and families, and perhaps during support hours outside of session. Further, in the age of managed care, clinicians may have a tendency to report multiple practices in treatment, regardless of their match with the evidence base, due to feeling systemic pressure that 'more is better' and this pressure could translate to treatment targets as well (Higa-McMillan et al., 2011).

Clinician-endorsed targets were also compared to the categorical ratings for targets associated with participants' diagnostic profiles. As described above, there was very little overlap between the targets associated with participants' diagnostic profiles, to which clinicians had access prior to beginning treatment, and the targets endorsed by clinicians (20%). The low level of consistency between diagnostically-related targets and clinician ratings may have resulted from a clinician focusing on a conceptually-related treatment target to the diagnoses, instead of the directly related categories created by the coding team. For example, for a participant with a diagnosis of Major Depressive Disorder, a clinician may have conceptualized the focus of treatment as "Contentment, Enjoyment, Happiness" or "Low Self Esteem," rather than explicitly "Depressed Mood." The current study did not examine this possibility; however, this is a direction for future research.

Potential Implications

The findings of this study support, in part, the utility of the Monthly Treatment Progress Summary (MTPS) as a measure of specific targets on which clinicians focus in treatment. Trained coders reliably identified specific treatment targets focused on in treatment sessions for 12 of 13 targets that occurred at a high enough frequency to be analyzed. However, it is important to highlight that these targets represent a small subset of the complete set of available targets that clinicians may endorse on the MTPS.

The discrepancy between the coding team's reliability across targets and the limited consistency between coders and clinicians provide valuable insight into the accuracy of clinicians' self-reported treatment content and the challenges in interpreting treatment target information strictly based on verbal exchanges during therapy sessions. On average, clinicians reported slightly more treatment targets, relative to the targets observed by a team of trained coders listening to audiotapes of treatment sessions. This has potential implications for service delivery (Hurlburt, Garland, Nguyen, & Brookman-Frazee, 2010). For example, it is possible that if a clinician overreports the targets on which there was a focus in session, such that targets that were not focused on or were not given sufficient focus, this will result in less effective treatment planning and delivery of targeted interventions.

Relatedly, although community clinicians provide services for many complex clients on their caseloads (Garland et al., 2010), and the diagnostic complexities of youth participants in the current study was no exception, streamlining treatment focus may be advantageous to reduce overwhelming clients/families and to be able to effectively assess incremental progress. Existing literature suggests that UC clinicians' treatment practices may be too varied or comprised of necessary and unnecessary components resulting in inefficient treatment plans and variable outcomes (Borntrager, Chorpita, Higa-McMillan, Daleidan, & Starace, 2013; Garland, Hawley,

Brookman-Frazee, & Hurlburt, 2008). The findings of this study - that clinicians tend to report more targets in session than coders detected - may indicate that clinicians' conceptualization of the treatment plan is too broad or varied and may result in less efficient services. Alternatively, clinicians may be targeting foci of treatment about which they were not explicit. Though this hypothesis underscores the measurement challenges of treatment targets (without complementing documentation such as treatment plans with explicit goals and objectives), at the same time it begs the question of whether or not the client is aware and engaged in their own treatment progression. That is, if trained coders were unable to detect certain targets of treatment, do the clients understand why they are in treatment or for what they are being treated? Importantly, this hypoethsis is spectulative and one to consider for future research given the limitations of the data collected, which is a small number of therapy sessions from a relatively discrete time in the overall course of a treatment episode.

Finally, given the increased demand for accountability within mental healthcare settings, the need for valid and reliable measures of mental health services is high (Glasgow & Riley, 2013). There are several practical and economic advantages to the use of an indirect measure, namely the lessening of burdens associated with valuable resources such as time, finances, and staffing. This study provides some support for the MTPS as a valid and reliable indirect measure of therapeutic techniques and content. Thus, the feasibility of the MTPS as a measure that is both economical advantageous, as well as reliable, is high. However, the low level of consistency between coders and clinicians indicate there is likely a need for the development and use of clearer instructions, treatment target definitions, and training for clinicians retrospectively reporting on the foci of their sessions.

During the coding process, coders noted that several treatment targets were difficult to interpret into observable behavior based on their definitions. For example, the treatment target *Positive Family Functioning* is defined as "issues related with healthy communication, problem-solving, shared pleasurable activities, physical and emotional support, etc. in the context of an interaction among multiple persons in a family relation, broadly defined." Following this definition, it was not clear whether interactions between a client and a clinician that focused on healthy communication with a member of the clients' family, who was not present in the session, fulfilled the definition of this target, or whether the family member being discussed needed to be present in session. As a coding team, we reached a decision to code by; however, this illustrates that there is an amount of leeway between the established definition of a treatment target and how that target is translated to session content. Similarly, for some treatment targets, the definitions seemed to overlap. For example, the definitions for the targets *Anxiety* and *Phobias/Fears* were quite similar, such that it was difficult to ascertain whether session content fell within one category over the other. Both of these treatment targets focused on anxiety, with the difference being that one referred to a more general state of anxiety and the other on specific situations, objects, or activities. In practical application, the line between when a treatment session was focused on a general experience of anxiousness or a specific fear or phobia became difficult to define. In order to facilitate more accurate reporting of treatment targets, as defined on the MTPS, it is likely that the measure would benefit from additional clarity in definitions and more comprehensive guidelines on the criteria differentiating similar targets.

Weersing et al. (2002) highlighted the issue that therapy is composed of both an observable action and underlying clinician intent. A strength of this study is in the use of both direct (audio-recorded and coded treatment sessions) and indirect (clinician self-reports)

measures from which we obtain observational information about clinician behavior as well as the clinicians' interpretation of their behavior. Despite this advantage, the low level of consistency observed between coders and clinicians suggest that the clinician self-report measure may tap into clinician intent, but that such intent is difficult to interpret through observable action (audio recorded treatment sessions). In other words, clinicians may endorse treatment targets that were the underlying intended focus of treatment, but that were not observed in the recorded treatment sessions.

The training that clinicians in this study received for the MTPS was conducted to reflect how clinicians are typically trained on the MTPS within the Child and Adolescent Mental Health Division (CAMHD). However, it is possible that the discrepancy observed between coders and clinicians resulted from clinicians and coders differing in their evaluation of what constituted an appropriate level of focus to report on the MTPS. Thus, it is possible that clinicians' intent influenced their ratings to reflect their intentions in sessions, and not the observable content on which the session was focused.

Limitations

This study has several limitations. First, the treatment data for this study was comprised of one month of treatment, which limited the number of targets that were available for analysis. Second, the coding process for the current study focused on targets that loaded onto the factor structure identified by Love and colleagues (2016), and not on the full list of targets available to clinicians in order to facilitate the feasibility of the project. However, by limiting the focus of this study, the observations are similarly constrained. The current study does not represent the full picture of clinician reporting patterns of treatment targets in this sample. It is possible that clinicians in this sample reported targets outside the model, utilizing treatment targets that were

not coded for in the present study to represent their treatment. Thus, it is possible that the observed level of consistency between coders and clinicians could have had a different result if all targets were included.

The current study utilized audio tapes of therapy sessions, which limited the scope of what coders could interpret as focus on a treatment target to audible verbal content. It is possible that there was session content rendered unobservable to the coders, and thus not coded. Clinicians may have been privy to a variety of visual cues or behaviors that influenced their reporting of treatment targets that coders were unable to examine. Relatedly, in the present study, the number of sessions available for each client differed (Table 3) as well as the length of sessions. Having more sessions available may have increased the likelihood that coders were able to deduce and code for a treatment target that matched the clinician self-reports.

Additionally, although the current study utilized both direct and indirect measures of treatment, it did not examine potential interactions between clinician intent and observable behavior. It is possible that the targets clinicians endorsed on the MTPS reflected their intentions for session, and not the actual content on which the session focused.

Coders may also have had a different perspective when coding for treatment targets. Coders were able to listen to treatment audio tapes multiple times if needed, without feedback from other coders, to develop their coding sheet. Thus, coders likely had the specifics of each treatment session more accessible than what clinicians had available as they completed the MTPS forms. In order to make complicated coding decisions, as described above, coders could rely on going back to the audio tape, where clinicians may have relied on their recollection of the events.

Finally, the coding process developed to endorse treatment focus utilized direct methods of observation, which may not fully represent the true intent of a clinician's behavior. For example, a clinician may have a particular target or focus in mind when asking a client a question or engaging in discussion that is not outwardly observable through the content of their speech. Not being fully privy to the internal iterations of clinicians' treatment intent may have contributed to the low agreement between clinicians and coders on targets of treatment.

Future Directions

This study sought to provide support for the validity of the MTPS as an indirect measure of treatment focus. For future studies, it will be important to address the shortcomings of the current study, and to expand upon its findings. Additional analyses should be conducted to examine the data regarding clinicians' patterns of target endorsement and whether a less direct relationship to participant diagnoses exists.

The current study's coding and validation procedures should be utilized in a sample of treatment data in which the remaining 35 targets are likely to be focused. Relatedly, the relationship between clinicians' endorsement of treatment targets and their endorsement of practice elements should be explored in a future study to determine potential interactions more formally. The results of the current study also suggest that providing more guidance or a more intensive training regarding reporting treatment targets on the MTPS could be useful, in addition to further development of the MTPS manual to clarify the differences between conceptually similar targets, and to provide more specification on the criteria that may include or exclude session content as applicable to a treatment target. Lastly, periodic quality control checks may be useful to maximize clinicians' valid, reliable responding.

Conclusion

The current study demonstrated some support for the use of the MTPS as an indirect measure of specific content on which clinicians focus in therapy. Trained coders reliably identified specific treatment targets on which treatment was focused for 12 of 13 targets that occurred at a high enough frequency to be analyzed. Overall, coder-clinician agreement was low, with four of the 12 targets achieving acceptable levels of reliability (ICC ≥ .60). These results suggest there may be a difference between clinician intent and the observable content in which a clinician engages. These results may also indicate differential levels of familiarity or understanding between coders and clinicians of what constituted focus in session, and provide an area of future development for the MTPS. The present study demonstrates that several treatment targets on the MTPS can be reliably coded, and it is likely that future research that incorporates a larger sample of more diverse content will indicate that many of the remaining treatment targets can also be reliably coded. The discrepancies noted between coders and clinicians indicate the need for future research to elucidate clinician intent, clinicians' accuracy in reporting their session content, and training implications.

References

Aarons, G. A., Hurlburt, M., & Horwitz, S. M. (2011). Advancing a conceptual model of evidence-based practice implementation in public service sectors. *Administration and Policy in Mental Health, 38*(1), 4-23.

Baker-Ericzén, M. J., Hurlburt, M. S., Brookman-Frazee, L., Jenkins, M. M., & Hough, R. L. (2010). Comparing child, parent, and family characteristics in usual care and empirically supported treatment research samples for children with disruptive behavior disorders. *Journal of Emotional and Behavioral Disorders, 18*(2), 82-99. doi:10.1177/1063426609336956

Balas, E. A., & Boren, S. A. (2000). Managing clinical knowledge for health care improvement. *Yearbook of Medical Informatics,* 65-70.

Bickman, L. (2000). The most dangerous and difficult question in mental health services research. *Mental Health Services Research, 2*(2), 71-72.

Bickman, L. (2008a). A measurement feedback system (MFS) is necessary to improve mental health outcomes. *American Academy of Child and Adolescent Psychiatry. Journal, 47*(10), 1114-1119.

Bickman, L. (2008b). Why don't we have effective mental health services? *Administration and Policy in Mental Health, 35*(6), 437-439.

Borntrager, C., Chorpita, B. F., Higa-McMillan, C. K., Daleiden, E. L., & Starace, N. (2013). Usual care for trauma-exposed youth: Are clinician-reported therapy techniques evidence-based? *Children and Youth Services Review, 35*(1), 133-141. doi:10.1016/j.childyouth.2012.09.018

Borntrager, C. F., Chorpita, B. F., Higa-McMillan, C., & Weisz, J. R. (2009). Provider attitudes toward evidence-based practices: Are the concerns with the evidence or with the manuals? *Psychiatric Services, 60*(5), 677-681.

Borntrager, C. F., Chorpita, B. F., Orimoto, T., Love, A., & Mueller, C. W. (2013). Validity of clinician's self-reported practice elements on the monthly treatment and progress summary. *The Journal of Behavioral Health Services & Research, 41*(1), 1-15.

Child and Adolescent Mental Health Division. (2003). *Service provider monthly treatment and progress summary*. Honolulu, HI: Hawaii Department of Health Child and Adolescent Mental Health Division.

Child and Adolescent Mental Health Division. (2008a). Instructions and codebook for service provider monthly treatment and progress summary. Retrieved from http://hawaii.gov/health/mental-health/camhd/library/pdf/paf/paf-001.pdf

Child and Adolescent Mental Health Division. (2008b). *Service provider monthly treatment and progress summary* Hawaii Department of Health Child and Adolescent Mental Health Division. Retrieved from http://hawaii.gov/health/mental-health/camhd/library/pdf/paf/paf-002.pdf

Chorpita, B. F., & Donkervoet, C. (2005). Implementation of the felix consent decree in hawaii. In R. G. Steele, & M. C. Roberts (Eds.), *Handbook of mental health services for children, adolescents, and families* (pp. 317-332). New York: Kluwer Academic/Plenum Publishers.

Chorpita, B. F., Weisz, J. R., Daleiden, E. L., Schoenwald, S. K., Palinkas, L. A., Miranda, J., . . . Gibbons, R. D. (2013). Long-term outcomes for the child STEPs randomized effectiveness trial: A comparison of modular and standard treatment designs with usual care. *Journal of Consulting and Clinical Psychology, 81*(6), 999-1009. doi:10.1037/a0034200

Cicchetti, D. V., & Sparrow, S. A. (1981). Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *American Journal of Mental Deficiency*, *86*(2), 127-137.

Clark, L. A., Watson, D., & Reynolds, S. (1995). Diagnosis and classification of psychopathology: Challenges to the current system and future. *Annual Review of Psychology, 46*(1), 121-153.

Daleiden, E., & Chorpita, B. F. (2005). From data to wisdom: Quality improvement strategies supporting large-scale implementation of evidence-based services. *Child and Adolescent Psychiatric Clinics of North America, 14*(2), 329-349.

Daleiden, E., Lee, J., & Tolman, R. (2004). *Annual evaluation report: Fiscal year 2004.* Honolulu, HI: Hawaii Department of Health Child and Adolescent Mental Health Division. Retrieved from http://hawaii.gov/health/mental-health/camhd/library/pdf/rpteval/ge/ge011.pdf

Elkin, I., Falconnier, L., Martinovich, Z., & Mahoney, C. (2006). Therapist effects in the national institute of mental health treatment of depression collaborative research program. *Psychotherapy Research, 16*(2), 144-160.

Fixsen, D. L., Naoom, S. F., Blase, K. A., & Friedman, R. M. (2005). *Implementation research: A synthesis of the literature.* Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute.

Garland, A. F., Bickman, L., & Chorpita, B. F. (2010). Change what? Identifying quality improvement targets by investigating usual mental health care. *Administration and Policy in Mental Health, 37*(1-2), 15-26.

Garland, A. F., Brookman-Frazee, L., Hurlburt, M. S., Accurso, E. C., Zoffness, R. J., Haine-Schlagel, R., & Ganger, W. (2010). Mental health care for children with disruptive behavior problems: A view inside therapists' offices. *Psychiatric Services, 61*(8), 788-795. doi: 10.1176%2Fappi.ps.61.8.788

Garland, A. F., Hawley, K. M., Brookman-Frazee, L., & Hurlburt, M. S. (2008). Identifying common elements of evidence-based psychosocial treatments for children's disruptive behavior problems. *Journal of the American Academy of Child & Adolescent Psychiatry, 47*(5), 505-514.

Garland, A. F., Hurlburt, M. S., Brookman-Frazee, L., Taylor, R. M., & Accurso, E. C. (2010). Methodological challenges of characterizing usual care psychotherapeutic practice. *Administration and Policy in Mental Health, 37*(3), 208-220.

Glasgow, R. E., & Riley, W. T. (2013). Pragmatic measures: What they are and why we need them. *American Journal of Preventive Medicine, 45*(2), 237–243. doi: 10.1016/j.amepre.2013.03.010

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology, 8*(1), 23-34.

Helzer, J. E., Kraemer, H. C., & Krueger, R. F. (2006). The feasibility and need for dimensional psychiatric diagnoses. *Psychological Medicine, 36*(12), 1671-1680.

Hembree, E., & Cahill, S. (2007). Obstacles to successful implementation of exposure therapy. In D. Richard, & D. Lauterbach (Eds.), *Handbook of exposure therapies* (pp. 389-408). Burlington, MA: Academic Press.

Higa-McMillan, C., Powell, C. K., Daleiden, E. L., & Mueller, C. W. (2011). Pursuing an evidence-based culture through contextualized feedback: Aligning youth outcomes and practices. *Professional Psychology: Research and Practice, 42*(2), 137-144.

Hurlburt, M. S., Garland, A. F., Nguyen, K., & Brookman-Frazee, L. (2010). Child and family therapy process: Concordance of therapist and observational perspectives. *Administration and Policy in Mental Health and Mental Health Services Research, 37*(3), 230-244.

Kelley, S. D., Vides de Andrade, Ana Regina, Sheffer, E., & Bickman, L. (2010). Exploring the black box: Measuring youth treatment process and progress in usual care. *Administration and Policy in Mental Health, 37,* 287-300. doi:10.1007/s10488-010-0298-8

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry, 62*(6), 617-627.

Kim, D., Wampold, B. E., & Bolt, D. M. (2006). Therapist effects in psychotherapy: A random-effects modeling of the national institute of mental health treatment of depression collaborative research program data. *Psychotherapy Research, 16*(2), 161-172.

Love, A., Okado, I., Orimoto, T., & Mueller, C. (2016). Factor analysis of therapist-identified treatment targets in community-based children's mental health. *Administration and Policy in Mental Health and Mental Health Services Research.* Advance online publication. doi:10.1007/s10488-016-0770-1

Love, A., Tolman, R., Mueller, C., & Powell, K. (2014). Frequency, level, and rate of improvement for treatment targets in a children's mental health community-based intensive in-home therapeutic setting. *Administration and Policy in Mental Health, 41*, 421–433. doi:10.1007/s10488-013-0480-x

McLeod, B. D., & Weisz, J. R. (2010). The therapy process observational coding system for child psychotherapy strategies scale. *Journal of Clinical Child and Adolescent Psychology, 39*(3), 436-443.

Nakamura, B. J., Daleiden, E. L., & Mueller, C. W. (2007). Validity of treatment target progress ratings as indicators of youth improvement. *Journal of Child and Family Studies, 16*(5), 729-741.

Nezu, A. M., & Nezu, C. M. (1993). Identifying and selecting target problems for clinical interventions: A problem-solving model. *Psychological Assessment, 5*(3), 254-263.

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

Southam-Gerow, M. A., Chorpita, B. F., Miller, L. M., & Gleacher, A. A. (2008). Are children with anxiety disorders privately referred to a university clinic like those referred from the public mental health system? *Administration and Policy in Mental Health, 35*(3), 168-180.

Verheul, R., & Widiger, T. A. (2004). A meta-analysis of the prevalence and usage of the personality disorder not otherwise specified (PDNOS) diagnosis. *Journal of Personality Disorders, 18*(4), 309-319.

Wampold, B. E., & Brown, G. S. (2005). Estimating variability in outcomes attributable to therapists: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology, 73*(5), 914-923. doi:10.1037/0022-006X.73.5.914

Weersing, V. R., Iyengar, S., Kolko, D. J., Birmaher, B., & Brent, D. A. (2006). Effectiveness of cognitive-behavioral therapy for adolescent depression: A benchmarking investigation. *Behavior Therapy, 37*(1), 36-48.

Weersing, V. R., Weisz, J. R., & Donenberg, G. R. (2002). Development of the therapy procedures checklist: A therapist-report measure of technique use in child and adolescent treatment. *Journal of Clinical Child and Adolescent Psychology: The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53, 31*(2), 168-180.

Weisz, J. R., Chorpita, B. F., Palinkas, L. A., Schoenwald, S. K., Miranda, J., Bearman, S. K., . . . Martin, 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., Jensen-Doss, A., & Hawley, K. M. (2006). Evidence-based youth psychotherapies versus usual clinical care: A meta-analysis of direct comparisons. *The American Psychologist, 61*(7), 671-689.

Widiger, T. A., & Samuel, D. B. (2005). Diagnostic categories or dimensions? A question for the diagnostic and statistical manual of mental disorders--fifth edition. *Journal of Abnormal Psychology, 114*(4), 494-504.

Zoffness, R., Garland, A., Brookman-Frazee, L., & Roesch, S. (2009). Case management as a significant component of usual care psychotherapy for youth with disruptive behavior problems. *Child and Youth Care Forum, 38*(4), 185-200.

Table 1

Characteristics of clinician participants in Child STEPs study

Characteristic	*N*	*M*
Age	15	38.3
Gender		
Female	10	
	(66.67)	
Male	5	
	(33.34)	
Ethnicity		
Asian American	6	
	(40)	
White	9	
	(60)	
Highest Degree Obtained		
Doctorate	2	
	(13.3)	
Masters	13	
	(86.7)	

Note. Values enclosed in parentheses are percentages.

Table 2

Characteristics of child participants in Child STEPs study

Characteristic	*N*	*M*
Age	19	9.55
Gender		
Female	4	
	(21)	
Male	15	
	(79)	
Ethnicity		
African American	1	
	(5)	
Asian American	1	
	(5)	
Mixed Race	10	
	(55)	
White	7	
	(35)	
Number of Sessions		2.7

Note. Values enclosed in parentheses are percentages.

Table 3

Number of Treatment Sessions per Client

Client ID	Number of Sessions
1	1
2	3
3	3
4	4
5	1
6	2
7	1
8	1
9	1
10	4
11	3
12	4
13	4
14	5
15	1
16	2
17	5
18	3
19	1

Table 4

Factor Names and Associated Treatment Targets

Factor Name	Treatment Target
Disinhibition	Oppositionality Aggression Hyperactivity Anger Empathy Attention Problems
Societal Rules Evasion	Willful Misconduct School Refusal/Truancy Substance Use Runaway
Social Engagement Deficits	Assertiveness Positive Thinking Social Skills Peer Involvement Phobias/Fears
Emotional Distress	Depressed Mood Suicidality Traumatic Stress Low Self-Esteem Grief
Management of Biodevelopmental Outcomes	Health Management Learning Disorder Positive Family Functioning Medical Regimen

Note. Adapted from "Factor Analysis of Therapist-Identified Treatment Targets in Community-Based Children's Mental Health", by A. Love, I. Okado, T. Orimoto, and C. Mueller, 2016, Advance online publication.

Table 5

Treatment Targets and Associated Diagnoses

Treatment Target	Associated Diagnoses
Anxiety	Any Anxiety Disorder (except Specific Phobia)
Attention Problems	ADHD-C or ADHD-PI
Cognitive/Intellectual Functioning	Any Cognitive Diagnosis
Depressed Mood	Any Mood Disorder or Adjustment Disorder with a mood component
Eating/Feeding Problems	Any Eating or Feeding Disorder
Enuresis/Encopresis	
Hyperactivity	ADHD-C or ADHD-PH
Mania	Bipolar Disorder
Oppositional/Non-Compliant Behavior	Oppositional Defiant Disorder
Phobia/Fears	Any Specific Phobia Disorder
Psychosis	Any Psychotic Disorder
Speech and Language Problems	Any Speech or Language Disorder
Substance Use	Any Substance Use Disorder
Traumatic Stress	Posttraumatic Stress Disorder, Acute Stress Disorder
Willful Misconduct/Delinquency	Conduct Disorder

Table 6

Frequency of Treatment Targets Endorsed by Coders and Clinicians per Client

Client ID	Number of Targets-Coders	Number of Targets-Clinicians
1	4	6
2	3	2
3	6	3
4	6	3
5	3	4
6	1	5
7	0	1
8	1	3
9	1	6
10	4	3
11	2	2
12	4	2
13	2	8
14	2	6
15	4	2
16	5	3
17	2	3
18	3	4
19	2	5

Table 7

Frequency of Treatment Targets Reported by Coders and Clinicians

Treatment Target	Coder-Frequency	Clinician-Frequency
Aggression	2	2
Anger	8	6
Assertiveness	1	2
Attention Problems	3	3
Depressed Mood	4	4
Grief	2	0
Hyperactivity	2	0
Oppositionality	8	6
Phobias/Fears	9	6
Positive Family Functioning	3	3
Positive Thinking	10	8
School Refusal	1	0
Social Skills	2	10

Table 8

Comparison of Coder-Coder Treatment Targets and Coder-Clinician Treatment Targets

Treatment Target	Coder-Coder ICC	Coder-Clinician ICC
Aggression	0.61	0.61
Anger	0.89	0.72
Assertiveness	0.89	0.00
Attention Problems	0.76	-0.46
Depressed Mood	0.62	0.23
Grief	0.79	0.00
Hyperactivity	0.59	---
Oppositionality	0.76	0.72
Phobias/Fears	0.70	0.77
Positive Family Functioning	0.60	0.00
Positive Thinking	0.79	0.29
School Refusal	0.95	0.00
Social Skills	0.80	0.49

Table 9

Comparison of Clinician Reported Treatment Targets and Categorical Diagnoses

Treatment Target	Categorical-Clinician ICC
Anxiety	0.88
Attention Problems	0.06
Cognitive-Intellectual Functioning	0.00
Depressed Mood	0.56
Eating, Feeding Problems	0.79
Hyperactivity	0.00
Oppositionality	-0.07
Phobias/Fears	0.34
Traumatic Stress	0.00
Willful Misconduct	0.00

Appendix A

Monthly Treatment Progress Summary

SERVICE PROVIDER MONTHLY TREATMENT & PROGRESS SUMMARY
Child and Adolescent Mental Health Division (CAMHD)

Instructions: Please complete and electronically submit this form to CAMHD by the 5th working day of each month (summarizing the time period of 1st to the last day of the previous month). The information will be used in service review, monitoring, planning and coordination in accordance with CAMHD policies and standards. Mahalo!

Client Name:	CR #:	DOB:
Month/Year of Services:	Eligibility Status:	Level of Care (one per form):
Axis I Primary Diagnosis:	Axis I Secondary Diagnosis:	Axis I Tertiary Diagnosis:
Axis II Primary Diagnosis:	Axis II Secondary Diagnosis:	

Service Format (circle all that apply):
Individual Group Parent Family Teacher Other: _____

Service Setting (circle all that apply):
Home School Community Out of Home Clinic/Office Other: _____

Service Dates:																

Targets Addressed This Month (*number* up to 10):

	Activity Involvement		Community Involvement		Hyperactivity		Positive Peer Interaction		Shyness
	Academic Achievement		Contentment, Enjoyment, Happiness		Learning Disorder, Underachievement		Phobia/Fears		Sleep Disturbance
	Adaptive Behavior/Living Skills		Depressed Mood		Low Self-Esteem		Positive Thinking/ Attitude		Social Skills
	Adjustment to Change		Eating, Feeding Problems		Mania		Pregnancy Education/ Adjustment		Speech and Language Problems
	Aggression		Empathy		Medical Regimen Adherence		Psychosis		Substance Use
	Anger		Enuresis, Encopresis		Occupational Functioning/Stress		Runaway		Suicidality
	Anxiety		Fire Setting		Oppositional/ Non-Compliant Behavior		School Involvement		Traumatic Stress
	Assertiveness		Gender Identity Problems		Peer Involvement		School Refusal/Truancy		Treatment Engagement
	Attention Problems		Grief		Peer/Sibling Conflict		Self-Control		Willful Misconduct, Delinquency
	Avoidance		Health Management		Personal Hygiene		Self-Injurious Behavior		Other:
	Cognitive-Intellectual Functioning		Housing/Living Situation		Positive Family Functioning		Sexual Misconduct		Other:

CR # _____ (please repeat the number here)

Progress Ratings This Month (check appropriate rating for any target numbers endorsed as targets):

#	Deterioration < 0%	No Significant Changes 0%-10%	Minimal Improvement 11%-30%	Some Improvement 31%-50%	Moderate Improvement 51%-70%	Significant Improvement 71%-90%	Complete Improvement 91%-100%	Date (If Complete)
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								

Intervention Strategies Used This Month (check all that apply):

	Activity Scheduling		Emotional Processing		Line of Sight Supervision		Personal Safety Skills		Stimulus or Antecedent Control
	Assertiveness Training		Exposure		Maintenance or Relapse Prevention		Physical Exercise		Supportive Listening
	Attending		Eye Movement, Tapping		Marital Therapy		Play Therapy		Tangible Rewards
	Behavioral Contracting		Family Engagement		Medication/ Pharmacotherapy		Problem Solving		Therapist Praise/Rewards
	Biofeedback, Neurofeedback		Family Therapy		Mentoring		Psychoeducation, Child		Thought Field Therapy
	Care Coordination		Free Association		Milieu Therapy		Psychoeducation, Parent		Time Out
	Catharsis		Functional Analysis		Mindfulness		Relationship or Rapport Building		Twelve-Step Program
	Cognitive		Goal Setting		Modeling		Relaxation		Other:
	Commands		Guided Imagery		Motivational Interviewing		Response Cost		Other:
	Communication Skills		Hypnosis		Natural and Logical Consequences		Response Prevention		Other:
	Crisis Management		Ignoring/Differential Reinforcement of Other Behavior		Parent Coping		Self-Monitoring		
	Cultural Training		Individual Therapy for Caregiver		Parent/Teacher Monitoring		Self-Reward/ Self-Praise		
	Discrete Trial Training		Insight Building		Parent/Teacher Praise		Skill Building		
	Educational Support		Interpretation		Peer Pairing		Social Skills Training		

CR # _____ (please repeat the number here)

Psychiatric Medications (List All)	Total Daily Dose	Dose Schedule	Check if Change	Description of Change
			☐	
			☐	
			☐	
			☐	
			☐	

Projected Discharge Date: _____ ☐ Check if Discharged During Current Month

IF YOUTH WAS DISCHARGED THIS MONTH, PLEASE COMPLETE ITEMS A & B:

A. Discharge Living Situation (check one):

☐ Home ☐ Foster Home ☐ Group Care ☐ Residential Treatment
☐ Institution/Hospital ☐ Jail/Correctional Facility ☐ Homeless/Shelter ☐ Other:_____

B. Reason(s) for Discharge (check all that apply):

☐ Success/Goals Met ☐ Insufficient Progress ☐ Family Relocation
☐ Runaway/Elopement ☐ Refuse/Withdraw ☐ Eligibility Change ☐ Other:_____

Outcome Measures: Optional. If you have any of the following data, please report the most recent scores:

CAFAS (8 Scales): (1-School:) (2-Home:) (3-Community:) (4-Behavior Toward Others:) (5-Moods/Emotions:) (6-Self-Harm:) (7-Substance:) (8-Thinking:) (Total:)			Date:
CASII/CALOCUS (Total):	CASII/CALOCUS (Level of Care):		Date:
CBCL (Total Problems T):	CBCL (Internalizing T):	CBCL (Externalizing T):	Date:
YSR (Total Problems T):	YSR (Internalizing T):	YSR (Externalizing T):	Date:
TRF (Total Problems T):	TRF (Internalizing T):	TRF (Externalizing T):	Date:
Arrested During Month? (Y/N):	School attendance (% of days):		

Comments/Suggestions (***attach additional sheets if necessary***):

Provider Agency & Island: _____	Clinician Name and ID#:_____
Provider Supervisor Signature: _____	Clinician Signature: _____
Submitted to CAMHD (date): _____	Care Coordinator: _____

Appendix B

Coding Manual

MTPS Treatment Target Validation Study
Coding Manual

Thank you for helping to push this study along! Below are the step-by-step instructions for coding for treatment targets. For this coding system, we will be combining micro- and macro-analytic coding techniques and we will be coding the 5 factors that demonstrated adequate loadings in Allison's factor analysis.

For example, we will be including a rating system that is continuous and rates how much of a focus a target was given on a micro-analytic scale (e.g., any TT coded in 5 minute increments). We will also be coding whether or not a target was explicitly found in a participant's consensus diagnostic profile.

CODING INSTRUCTIONS:
Before you start listening to a session, be sure to indicate the session date (NOT today's date). Also, indicate the client number in the blank. All TTs are coded within 5-minute increments according to the following scale:

CATEGORICAL RATING
0 = no factor or TT was identified
1 = a factor or TT was identified in the diagnostic profile

TTs = 1 matching diagnostic profile (composite):

-Anxiety (ANY ANXIETY DX, INCLUDING OCD, NOT SPECIFIC PHOBIA)
-Attention problems (ADHD-C OR ADHD-PI)
-Cognitive/intellectual functioning (ANY COGNITIVE DX)
-Depressed mood (ANY MOOD DX OR ADJUSTMENT WITH MOOD COMPONENT)
-Eating/feeding problems (ANY EATING OR FEEDING DISORDER, INCLUDING PICA)
-Enuresis/encopresis
-Hyperactivity (ADHD-C OR ADHD-HI)
-Mania (BIPOLAR DX)
-Oppositional/non-compliant behavior (ODD)
-Phobia/fears (ANY SPECIFIC PHOBIA)
-Psychosis (ANY PSYCHOTIC DX)
-Speech and language problems (ANY SPEECH/LANGUAGE DX)
-Substance use (ANY SUBSTANCE USE DX)
-Traumatic stress (PTSD; ACUTE STRESS DX)
-Willful misconduct/delinquency (CONDUCT DX)

FOCUS RATING
0 = no implicit or explicit focus or mention of the factor whatsoever
1 = a fleeting therapist action or statement that implies a factor
2 = at minimum, explicit discussion or activity clearly targeting a specific factor by name/synonym/TT definition subsumed under a factor OR greater than or equal to 50% of TTs under a factor receiving a 1 or higher

Factor definitions:

Disinhibition – reflects treatment targets related to oppositionality, inattention, and interpersonal difficulties. *Six* targets on disinhibition are:

1. oppositionality: behaviors that can be described as refusal to follow adult requests or demands or established rules and procedures (e.g., classroom rules, school rules, etc.).
2. aggression: verbal and/or physical aggression, or threat thereof, that results in intimidation, physical harm, or property destruction.
3. hyperactivity: can be described by fidgeting, squirming in seat, inability to remain seated, talking excessively, difficulty engaging in leisure activities quietly, etc.
4. anger: emotional experience or expression of agitation or destructiveness directed at a particular object or individual. Common physical feelings include accelerated heartbeat, muscle tension, quicker breathing, and feeling hot.
5. empathy: identifications with and understanding of another person's situation, feelings, and motives.
6. attention problems: described by short attention span, difficulty sustaining attention on a consistent basis, and susceptible to distraction by extraneous stimuli.

Unlawful behaviors – severely problematic disruptive behaviors. *Four* targets on unlawful behaviors are:

1. willful misconduct: persistent failure to comply with rules or expectations in the home, school, or community. Excessive fighting, intimidation of others, cruelty or violence toward people or animals, and/or destruction of property.
2. school refusal/truancy: reluctance or refusal to attend school without adult permission for the absence. May be associated with school phobia or fear manifested by frequent somatic complaints associated with attending school or in anticipation of school attendance, or willful avoidance of school in the interest of pursuing other activities.
3. substance use: issues related to the use or misuse of a common, prescribed, or illicit substances for altering mental or emotional experience or functioning.
4. runaway: running away from home or current residential placement for a day or more.

Overcoming withdrawal - targets largely focused on enhancing interpersonal skills. *Five* targets on overcoming withdrawal are:

1. assertiveness: the skills or effectiveness of clearly communicating one's wishes. For example, the effectiveness with which a child refuses unreasonable requests from others, expresses his/her rights in a non-aggressive manner, and/or negotiates to get what s/he wants in their relationships with others.
2. positive thinking: this target involves clear, healthy, or optimistic thinking, and involves the absence of distortions or cognitive bias that might lead to maladaptive behavior.
3. social skills: skills for managing interpersonal interactions successfully. Can include body language, verbal tone, assertiveness, and listening skills, among other areas.

4. peer involvement: greater involvement in activities with peers. Activities could range from academic tasks to recreational activities while involvement could range from working next to a peer to initiating an activity with a peer.
5. phobias/fears: irrational dread, fear, and avoidance of an object, situation, or activity.

Distress coping – targets focused on managing distressing emotions. *Five* targets on distress coping are:

1. depressed mood: behaviors that can be described as persistent sadness, anxiety, or "empty" mood, feelings of hopelessness, guilt, worthlessness, helplessness, decreased energy, fatigue, etc.
2. suicidality: issues related to recurrent thoughts, gestures, or attempts to end one's life.
3. traumatic stress: issues related to recurrent thoughts, gestures, or attempts to end one's life.
4. low self-esteem: an inability to identify or accept his/her positive traits or talents, and accept compliments. Verbalization of self-disparaging remarks and viewing him or herself in a negative manner.
5. grief: feelings associated with a loss of contact with a significant person in the youth's environment (e.g., parent, guardian, friend, etc.).

Biodevelopmental issues – targets related to health management and learning disability. *Four* targets on biodevelopmental issues are:

1. health management: issues related to the improvement or management of one's health, inclusive of both physical illness and fitness. In addition to dealing with the general development of health-oriented behavior and management of health conditions, this target can also focus on exercise or lack of exercise.
2. learning disorder: Refers to specific challenges with learning or educational performance that are not better accounted for by cognitive-intellectual functioning (#9) or general academic achievement (#1).
3. positive family functioning: issues related with healthy communication, problem-solving, shared pleasurable activities, physical and emotional support, etc. in the context of an interaction among multiple persons in a family relation, broadly defined.
4. medical regimen: knowledge, attitudes, and behaviors related to regular implementation procedures prescribed by a health care professional. Commonly include lifestyle behaviors (e.g., exercise, nutrition), taking medication, or self-administration of routine assessments (e.g., taking blood samples in a diabetic regimen).

Appendix C

Coding Sheet

Session Date:										Client #:_____					Client #:_____
	MTPS Coding Sheet														
Time Interval	0-4.59	5-9.59	10-14.59	15-19.59	20-24.59	25-29.59	30-34.59	35-39.59	40-44.59	45-49.59	50-54.59	55-59.59	60-64.59	In Dx Profile?	ategorica	Notes
Factors																
Disinhibition CATEGORICAL																
Disinhibition FOCUS																
Oppositionality																
Aggression																
Hyperactivity																
Anger																
Empathy																
Attention Problems																
Unlawful behaviors CATEGORICAL																
Unlawful behaviors FOCUS																
Willful Misconduct																
School Refusal/Truancy																
Substance Use																
Runaway																
Overcoming withdrawal CATEGORICAL																
Overcoming withdrawal FOCUS																
Assertiveness																
Positive Thinking																
Social Skills																
Peer Involvement																
Phobias/Fears																
Distress coping CATEGORICAL																
Distress coping FOCUS																
Depressed Mood																
Suicidality																
Traumatic Stress																
Low Self-Esteem																
Grief																
Biodevelopmental issues CATEGORICAL																
Biodevelopmental issues FOCUS																
Health Management																
Learning Disorder																
Positive Family Functioning																
Medical Regimen																
Additional TTs for Dx profile																
Anxiety																
Cognitive/Intellectual functioning																
Eating/Feeding Problems																
Enuresis/Encopresis																
Mania																
Psychosis																
Speech/Language Problems																