



# Predicting Youth Improvement in Community-Based Residential Settings with Practices Derived from the Evidence-Base

Sonia C. Izmirian<sup>1</sup> · Jaime P. Chang<sup>1</sup> · Brad J. Nakamura<sup>1</sup>

Published online: 14 February 2019  
© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

The current investigation conducted descriptive analyses on key variables in community-based residential (CBR) settings and investigated the extent to which disruptive youth between the ages of 13 and 17 years improved based on therapists' reported alignment with using practices derived from the evidence-base (PDEBs). Results from both the descriptive analyses and multilevel modeling suggested that therapists are using practices that both do and do not align with the evidence-base for disruptive youth. In addition, both PDEBs and practices with minimal evidence-support predicted or marginally predicted final average progress rating for these youth. Findings are discussed as they relate to the importance of continued exploration of treatment outcomes for CBR youth.

**Keywords** Evidence-based practices · Practice elements · Community-based residential · Youth improvement

“It is difficult and perhaps foolhardy to try to improve what you do not understand” (Hoagwood and Kolko 2009, p. 35). Applied to youth mental health system improvement efforts, we may benefit greatly from systematic and scientific study of large-scale public sector service systems. Although a significant amount of money is spent on children's mental health services (Soni 2009), research on the use of therapy practices in treatment as usual settings is scarce and tends to be an unexamined “black box” (Bickman 2000; Hoagwood and Kolko 2009). More research is needed to understand youth mental health systems, therapeutic approaches that help youth improve, and factors that mediate and moderate quality improvement efforts (Weisz and Jensen 2001).

Treatment as usual research on therapeutic approaches carries several potential benefits (e.g., developing a common language, describing services offered, training clinicians, providing new quality improvement methods; Garland et al. 2006). However, the systematic and scientific study of treatment as usual has been difficult due to the lack of standardized reporting metrics for treatment strategies used by therapists (Bickman 2000). For example, evidence-based practices can be specified at the level of brand-name

treatment manuals, treatment families (e.g., cognitive behavior therapy), or practice elements. As part of Chorpita, Daleiden, and Weisz's Distillation and Matching Model (2005, p. 11), a practice element is defined as “a discrete clinical technique or strategy (e.g., “time out,” “relaxation”) used as part of a larger intervention plan (e.g., a manualized treatment program for youth depression).” Based on this definition, a treatment manual is conceptualized as being made up of several discrete practice elements.

Distilling large numbers of empirically-supported youth mental health treatment protocols into patterns of technique commonalities has proven fruitful for numerous quality improvement strategies, such as guiding the creation of modular treatment approaches for complex youth (Chorpita et al. 2005; Weisz et al. 2012), service system feedback efforts (Higa-McMillan et al. 2011), and therapist training efforts (Southam-Gerow et al. 2013). More importantly for the current study, defining evidence-based practices at the practice element level has helped researchers investigate discrete technique level implementation in treatment as usual settings (e.g., Garland et al. 2008; Love 2014; Orimoto et al. 2013), and build upon studies that previously characterized such work as highly eclectic in nature (Norcross and Karpiak 2012).

✉ Sonia C. Izmirian  
sizmirian@gmail.com

<sup>1</sup> University of Hawaii at Manoa, 2530 Dole Street, Sakamaki Hall C400, Honolulu, HI 96822, USA

## Measuring and Identifying Practice Elements Derived from the Evidence-Base

When investigating practice element treatment patterns and associations within treatment as usual settings, researchers sometimes focus on a subset of practice elements that have been commonly referred to as practices derived from the evidence-base (PDEBs; Higa-McMillan et al. 2014). PDEBs are typically defined as practice elements found in a significant proportion of larger empirically-supported manualized approaches for any given problem area of interest (Higa-McMillan et al. 2014). For example, the practice element of exposure is considered a PDEB for anxiety given that it is present across 86% of study groups for all anxiety empirically supported approaches.

One method of identifying PDEBs in this fashion that has been steadily gaining momentum since 2009 (e.g., Mueller et al. 2009; Southam-Gerow et al. 2013) is through the PracticeWise, LLC Evidence-Based Youth Mental Health Services coding system. Within this system, the treatment outcome literature is routinely coded and summarized along both efficacy (e.g., practice element aggregate summaries across empirically supported treatment manuals for a problem area) and contextual (e.g., gender, age, ethnicity, and settings) parameters. Treatment support for these manuals is also defined along five Levels ranging from “Best Support” to “No Support” (Chorpita and Daleiden 2009). The field’s youth mental health treatment outcome literature can be summarized and aggregated through various reports via the Evidence-Based Youth Mental Health Services Literature Database search engine (PracticeWise, LLC 2014). When distilled across empirically-supported treatment approaches (usually defined as Level-Two support or better) and specific parameters (e.g., male, age 15, residential setting), practice elements are said to aggregate into “profiles” (Izmirian et al. 2016) noting the percent of empirically supported treatment groups containing any given practice element (PracticeWise, LLC 2014).

## Previous Treatment as Usual Research on Therapeutic Practices

Previous treatment as usual research has investigated the associations between practice element utilization patterns and therapist characteristics, family characteristics, and youth outcomes. Regarding therapist information, studies have indicated that more PDEB-use in treatment as usual settings was related to a cognitive-behavioral theoretical

orientation and less experience (Brookman-Fraze et al. 2010), favorable attitudes toward evidence-based practices (Baumann et al. 2006), and psychology/psychiatry specialty (Higa-McMillan et al. 2014). Regarding family characteristics, studies performed specifically on youth with disruptive behavior problems have reported that more evidence-based practices were used with older youth, youth with more severe symptoms, and caregivers with higher education levels (Brookman-Fraze et al. 2010). Furthermore, the array and dosages of practice elements used were found to vary based on the number and types of youth diagnoses (Garland et al. 2010; Orimoto et al. 2014). For example, one study of 519 youth with various problem areas in Hawai’i found that PDEB-use was predicted by longer lengths of treatment, while the use of practices with minimal evidence support (PMES) was predicted by older age, males, out-of-home levels of care, and youth not receiving evidence-based programs (Higa-McMillan et al. 2014). Taken together, this group of studies suggests that PDEB and practice element utilization patterns can be associated with a wide variety of therapist and youth characteristics, and that no unifying or coherent picture has yet emerged.

An influential area of treatment as usual research concerns the relationship between practice element utilization patterns and youth improvement. Studies with depressed youth suggested that higher end of treatment improvement ratings were predicted by more use of cognitive behavior therapy (as compared to treatment as usual; Weersing and Weisz 2002) and by more PDEBs, lower functional impairment, and lower number of diagnoses (Love 2014). Similarly with externalizing youth, higher rates of improvement were found if higher proportions or higher numbers of PDEBs (as opposed to PMESs) were used for each problem area examined (Denenny and Mueller 2012; Mueller et al. 2009; Orimoto et al. 2013).

A commonality across the majority of studies reviewed above is that they primarily focus on lower levels of care (e.g., in-home). The extent to which these types of results can generalize to the other levels of care remains open to empirical inquiry. Systematic study of youth placed into higher levels of care (e.g., out-of-home) is an essential next step in order to (a) help these youth improve and return to their homes, (b) prioritize the study of youth with higher functional impairment (Daleiden et al. 2010), and (c) possibly reduce the higher operating costs for out-of-home services (Lyons and McCulloch 2006).

## Community-Based Residential Services

Although it is arguably important to study all out-of-home treatment settings, the population of focus for this study was youth placed into community-based residential (CBR)

settings because of their longer treatment length (Orimoto et al. 2012) and higher referral rates as compared to other out-of-home settings (Keir et al. 2014). The current investigation focused on the improvement rate of disruptive youth within the CBR setting due to this problem area being the most referred to public mental health systems and CBR settings (Keir et al. 2014; Kazdin 1995), and being associated with the most substantial impact on society (e.g., cost, difficulties into adulthood; Pardini and Fite 2010). Given the sparse research in this area, the first aim of this study was to conduct descriptive analyses on key youth and therapist clinical factors to gain a better understanding of the youth typically served in CBR settings within a youth public mental health system. The second aim of this study was to use multilevel modeling to investigate the rate at which disruptive youth in the CBR setting improve based on therapists' reported use of PDEBs. It was hypothesized that higher rates of using PDEBs would be associated with greater rates of improvement on progress ratings for disruptive youth in the CBR setting.

## Methods

### System of Care

The State of Hawai'i's Department of Health's Child and Adolescent Mental Health Division provides youth mental health services for multiple levels of care through various contracted provider agencies. CBR<sup>1</sup> programs provide 24-hour residential treatment for behavioral, emotional, and/or family problems in supervised, safe, and therapeutic environments, and include onsite youth education, diagnostic, and therapy services aimed at providing social and daily living skills (Child and Adolescent Mental Health Division 2012). Five agencies were included in this study because they met the inclusion criteria (discussed below) and had more than 10 participants enrolled for their services during the study period. These agencies varied in specialization (e.g., substance use, males with conduct issues, females with emotional and behavioral issues) and operation dates.

<sup>1</sup> Within the Child and Adolescent Mental Health Division, only level III of the three levels of community-based residential (CBR) settings was included in the study because youth from the Community-Based Residential-I and -II settings have treatment needs that are very specialized (e.g., sexual deviance), as compared to the broad difficulties faced by CBR-III youth. Henceforth, the use of CBR will represent the youth from CBR-III programs only.

## Participants

### Youth Participants

Participants were included in the study if they were (a) between the ages of 13 to 17 (described in more detail in the "Data Analytic Strategy" section), (b) received services between 2006 and 2014, (c) completed at least two months of CBR treatment, (d) had at least one of seven disruptive behavior targets (discussed below) endorsed on the Monthly Treatment Progress Summary (MTPS; see Measures) for each month while placed in the CBR agency, and (e) had at least two unique disruptive behavior targets of the seven endorsed at any time within the youth's treatment episode to obtain a sample with diverse disruptive behavior problems (cf. Orimoto et al. 2013). Table 1 provides the demographic information for the 341 youth that met the inclusion criteria. Overall, the typical youth placed into CBR settings was a 15.8-year-old male, who was multiracial, had three diagnoses, stayed 6 months, and was discharged successfully to the home setting.

### Therapist Participants

Therapists at CBR programs have various specialties. Since typically there are multiple therapists working with a youth in CBR settings (Child and Adolescent Mental Health Division 2012), the therapist listed as the "clinician" on the MTPS form was considered the lead therapist<sup>2</sup> for each youth. Forty-seven unique therapists were associated with the sample of 341 youth. Archival therapist data included number of degrees, highest degree, professional specialty, licensure status, and credentialing position within the Child and Adolescent Mental Health Division (i.e., Qualified Mental Health Professional, Mental Health Professional, and Paraprofessional). Therapist demographic information in Table 2 suggested that the modal therapist was characterized as being an unlicensed Mental Health Professional, who had one master's degree in the specialty of marriage and family therapy or counseling psychology.

<sup>2</sup> When multiple therapists were tied to one youth's MTPS forms, the therapist that was most frequently linked to the MTPS forms was chosen for analyses. When multiple therapists were tied to one youth's MTPS forms and the therapists had the same number of MTPS forms completed, the initial therapist was chosen because previous research suggests that a youth typically improves at higher rates earlier as compared to later in treatment (Orimoto et al. 2012b).

**Table 1** Youth demographic and clinical information for the total sample size (n = 341)

Variable	Total sample
Sample size <sup>a</sup>	341 (100.0%)
Age	15.8 (1.2)
Gender (male) <sup>a</sup>	271 (79.5%)
Length of CBR-III Episode (days)	187.9 (82.5)
Race <sup>a</sup>	
Asian	21 (6.2%)
Multiracial	207 (60.7%)
Native Hawaiian or other Pacific Islander	61 (17.9%)
White	32 (9.4%)
Other	6 (1.8%)
Not available	14 (4.1%)
Any diagnosis present <sup>a</sup>	
Attention deficit/hyperactivity disorder	25 (7.3%)
Bipolar	9 (2.6%)
Depression	26 (7.6%)
Disruptive behavior	110 (32.3%)
Substance use	101 (29.6%)
Trauma	10 (2.9%)
Other	104 (30.5%)
Missing	199 (58.4%)
Number of diagnoses	3.0 (0.9)
Discharge situation <sup>a</sup>	
Foster home	24 (7.0%)
Home	152 (44.6%)
Homeless/shelter	12 (3.5%)
Institution/Hospital	6 (1.8%)
Jail/correctional facility	8 (2.3%)
Other	83 (24.3%)
Missing	56 (16.4%)
Discharge status <sup>ab</sup>	
Success	167 (49.0%)
Insufficient progress	22 (6.5%)
Runaway	73 (21.4%)
Refused treatment	20 (5.9%)
Other	26 (7.6%)
Missing	54 (15.8%)
CAFAS total	131.5 (27.6)
School role performance	26.1 (7.5)
Home role performance	26.9 (6.1)
Community role performance	20.0 (7.6)
Behavior toward others	19.8 (5.7)
Moods/emotion	16.0 (7.0)
Moods/self-harmful behavior	2.6 (6.6)
Substance use	17.6 (11.2)
Thinking	2.4 (5.7)

Any diagnosis present represents the percent of youth who had a diagnosis in each category anywhere on their diagnostic profile, regardless of order (primary, secondary, tertiary etc.). Anxiety and Adjustment were included in Other diagnoses due to their low frequency

**Table 1** (continued)

All other variables represent means and standard deviations. The sub-scale scores for the CAFAS can have a range from 0 to 30, with 0, 10, 20, and 30 representing minimal, mild, moderate, and severe impairment, respectively. The total score for the CAFAS can have a range from 0 to 240

CAFAS Child and Adolescent Functional Assessment Scale, CBR community-based residential

<sup>a</sup>Represents frequencies and percentages

<sup>b</sup>It should be noted that there were occasionally multiple discharge statuses marked off for one youth (e.g., runaway and refused treatment)

**Table 2** Therapist information for the total sample size for both clients (n = 341) and therapists (n = 47)

Variable	Total client sample (n = 341)	Total therapist sample (n = 47)
Number of degrees	1.0 (0.1)	1.0 (0.15)
Highest degree <sup>a</sup>		
High School Diploma or GED	1 (0.3%)	1 (2.1%)
Associate/vocational/certificate	2 (0.6%)	1 (2.1%)
Bachelors (BA, BS)	10 (2.9%)	4 (8.5%)
Masters (MA, MS, MSW, MFT)	249 (73.0%)	38 (80.9%)
Doctorate of Psychology (PsyD)	6 (1.8%)	1 (2.1%)
Medical Degree (MD)	70 (20.5%)	1 (2.1%)
Juris Doctor (JD)	3 (0.9%)	1 (2.1%)
Professional specialty <sup>a</sup>		
Clinical psychology	7 (2.1%)	2 (4.3%)
Counseling psychology	82 (24.0%)	11 (23.4%)
Marriage and family therapy	109 (31.9%)	12 (25.5%)
Law/political science	11 (3.2%)	3 (6.4%)
Social work	34 (10.0%)	12 (25.5%)
Substance abuse counseling	2 (0.6%)	1 (2.1%)
Other (due to High School)	1 (0.3%)	1 (2.1%)
Psychology	25 (7.3%)	4 (8.5%)
Medicine	70 (20.5%)	1 (2.1%)
Licensed <sup>a</sup>	108 (31.7%)	8 (17.0%)
Position <sup>a</sup>		
Qualified Mental Health Professional	99 (29.0%)	6 (12.8%)
Mental Health Professional	229 (67.2%)	35 (74.5%)
Paraprofessional	13 (3.8%)	6 (12.8%)

<sup>a</sup>Represents frequencies and percentages. All other variables represent means and standard deviations

## Measures

### Monthly Treatment Progress Summary (MTPS; Child and Adolescent Mental Health Division 2005)

The MTPS is a therapist-report form designed to collect information on service formats, settings, dates, treatment

targets, practice elements, progress ratings, medications, and discharge. State-contracted service providers are mandated to fill out the MTPS monthly for each of their clients, and data are routinely utilized for an array of ongoing quality improvement efforts (Nakamura et al. 2014). Therapists identify up to ten treatment targets addressed during each reporting month, from a list of 48 predefined responses and two write-in fields. Treatment targets have evidenced convergent validity with related diagnoses, divergent validity with unrelated diagnoses, moderate stability over time, and a five-factor structure (Daleiden et al. 2004; Love et al. 2011; Nakamura et al. 2007). Therapists then assign progress ratings to each of the identified targets, based on the degree of progress achieved between the child's baseline level of functioning and the goal for that target. Progress ratings are ranked on a seven-point Likert scale: 0 = Deterioration, 1 = No Significant Change, 2 = Minimal Improvement, 3 = Some Improvement, 4 = Moderate Improvement, 5 = Significant Improvement, and 6 = Complete Improvement. Progress ratings have been found to correlate significantly with other measures of functioning in expected directions (Nakamura et al. 2007). Therapists are also instructed to indicate intervention strategies (i.e., practice elements) utilized with the youth in the given MTPS month from a list of 63 predefined responses and three write-in options. Referred to as "intervention strategies" on the form itself, the list of 63 predefined practice elements represent a mix of PDEBs and PMESs, distilled from previous literature reviews and ongoing analyses of frequently occurring write-ins (Chorpita et al. 2005). The practice elements have demonstrated moderate stability over time, a three-factor structure with good internal reliability, and convergent validity between reported and observed behaviors (Daleiden et al. 2004; Borntreger et al. 2013; Orimoto et al. 2012a).

### Child and Adolescent Functional Assessment Scale (Hodges, 1994)

The Child and Adolescent Functional Assessment Scale is a 200-item scale that measures level of functional impairment. Therapists score youth on their highest level of impairment (i.e., severe = 30, moderate = 20, mild = 10, no/minimal = 0) based on specific items in each of the eight domains of functioning: School Role Performance, Home Role Performance, Community Role Performance, Behavior Toward Others, Mood/Emotions, Mood/Self-Harmful Behavior, Substance Use, and Thinking. The total Child and Adolescent Functional Assessment Scale score (range = 0 to 240) is calculated by summing across the eight domains. For the purposes of this study, a client's baseline Child and Adolescent Functional Assessment Scale score (i.e., within 45 days from the start date), completed by Child and Adolescent Mental Health Division Care Coordinators, was entered

as a covariate in the study at the client-level (c.f., Orimoto et al. 2013). The Child and Adolescent Functional Assessment Scale has evidenced acceptable internal consistency across items, adequate convergent validity with other related measures and metrics, good inter-rater reliability across different samples of raters, and predictive relationships with service utilization and costs of services (Hodges and Gust 1995; Hodges and Wong 1996; Keir et al. 2014; Mueller et al. 2010; Nakamura et al. 2007).

## Procedures

### Data Source and Human Subjects Considerations

The Child and Adolescent Mental Health Division's Research Evaluation and Training Program electronically extracted a limited dataset with relevant client and therapist clinical and demographic data from the Child and Adolescent Mental Health Division's data management systems, which is compliant with the Health Insurance Portability and Accountability Act and Family Educational Rights and Privacy Act standards. The youth's legal guardian signed the Notice of Privacy Practices consent form, which allowed for the use of data for research purposes. The University of Hawai'i at Mānoa's Institutional Review Board also approved this study.

### Data Analytic Strategy

#### Defining Treatment Targets, PDEB-, and PMES-Scores

The disruptive behavior problem area was defined by seven treatment targets utilized by PracticeWise, LLC's (2014) literature coding summaries and PDEB profiles for "disruptive behavior" (bolded in Table 3). Due to a limited number of youth mental health randomized controlled trials conducted in the CBR setting, a list of PDEBs were compiled for youth from all levels of care to provide a broader, less nuanced, list of practices (Love 2014; Orimoto et al. 2013). Additionally, although youth referred to the CBR setting are typically between the ages of 12 and 17 (Child and Adolescent Mental Health Division 2012), previous research has indicated that there are differences in evidence-based practice patterns for disruptive youth between the age ranges of 12 and below and 13 and above (Chorpita and Daleiden 2009). Hence, the practice element profile and sample for this study were limited to youth 13 to 17 years old, and PDEBs were distilled from the treatment outcome literature for disruptive youth, 13 years or older, from all levels of care. For the purposes of this study, practices were considered PDEB if they were present in at least 30% of study groups with Level-Two or better support (cf. Denenny and Mueller 2012; Love 2014). Based on this information, there were 15 practice elements

**Table 3** Frequency counts of treatment targets for the total sample (n = 341)

Treatment target (TT)	TT utilized at least once throughout the episode	TT utilization overall
<b>Oppositional/non-compliant behavior</b>	<b>286 (83.9%)</b>	<b>76.0% (37.6)</b>
Substance use	298 (87.4%)	69.0% (37.4)
Positive peer interaction	289 (84.8%)	68.1% (39.3)
<b>Anger</b>	<b>289 (84.8%)</b>	<b>67.5% (38.1)</b>
Activity involvement	275 (80.6%)	67.3% (41.3)
Phobia/fears	252 (73.9%)	56.7% (42.2)
Treatment engagement	242 (71.0%)	51.7% (41.9)
<b>Aggression</b>	<b>216 (63.3%)</b>	<b>44.7% (41.9)</b>
<b>Runaway</b>	<b>209 (61.3%)</b>	<b>40.8% (41.8)</b>
<b>Willful misconduct/delinquency</b>	<b>183 (53.7%)</b>	<b>34.4% (39.4)</b>
Avoidance	171 (50.1%)	30.6% (37.9)
Empathy	142 (41.6%)	23.8% (35.2)
Academic Achievement	144 (42.2%)	21.6% (32.6)
Self-injurious behavior	140 (41.1%)	21.3% (33.5)
Anxiety	127 (37.2%)	19.7% (32.6)
Low self-esteem	108 (31.7%)	19.6% (35.0)
Social skills	130 (38.1%)	18.4% (30.1)
Depressed mood	114 (33.4%)	18.1% (32.4)
Positive thinking/attitude	116 (34.0%)	16.6% (29.0)
Peer involvement	115 (33.7%)	14.4% (25.5)
Contentment/enjoyment/happiness	101 (29.6%)	13.0% (25.2)
Assertiveness	96 (28.2%)	12.7% (26.2)
Peer/sibling conflict	80 (23.5%)	10.2% (23.2)
Community involvement	81 (23.8%)	10.2% (22.1)
Adjustment to change	63 (18.5%)	8.6% (23.3)
School involvement	90 (26.4%)	8.6% (19.6)
Attention problems	68 (19.9%)	7.9% (20.1)
Medical regimen adherence	51 (15.0%)	7.7% (22.6)
Traumatic stress	38 (11.1%)	7.0% (23.5)
Hyperactivity	41 (12.0%)	6.7% (21.5)
Grief	30 (8.8%)	5.7% (21.2)
Self-control	36 (10.6%)	5.1% (18.6)
Other*	25 (7.3%)	4.4% (17.7)
School attendance	37 (10.9%)	3.9% (15.3)
Cognitive-intellectual functioning	35 (10.3%)	3.0% (12.5)
Adaptive behavior/living skills	22 (6.5%)	2.1% (10.6)
Health management	18 (5.3%)	1.8% (10.9)
Positive family functioning	9 (2.6%)	1.6% (11.2)
Suicidality	7 (2.1%)	1.4% (10.7)
Shyness	9 (2.6%)	1.1% (9.1)
<b>Sexual Misconduct</b>	<b>15 (4.4%)</b>	<b>1.1% (7.3)</b>
Mania	5 (1.5%)	0.7% (7.2)
Learning disorder/underachievement	9 (2.6%)	0.7% (5.3)
Sleep disturbance	4 (1.2%)	0.5% (4.8)
Eating/feeding problems	4 (1.2%)	0.5% (4.6)
Psychosis	3 (0.9%)	0.4% (3.9)
Gender identity problems	3 (0.9%)	0.3% (2.9)
Housing/living situation	7 (2.1%)	0.3% (2.2)
Personal hygiene	2 (0.6%)	0.2% (3.8)
Treatment planning*	1 (0.3%)	0.2% (3.4)

**Table 3** (continued)

Treatment target (TT)	TT utilized at least once throughout the episode	TT utilization overall
Occupational functioning/stress	4 (1.2%)	0.2% (1.5)
Parenting skills*	1 (0.3%)	0.1% (2.2)
Compulsive behavior*	1 (0.3%)	0.0% (0.7)
Unclear*	1 (0.3%)	0.0% (0.5)
Enuresis/encopresis	0	0
<b>Fire setting</b>	<b>0</b>	<b>0</b>
Pregnancy education/adjustment	0	0
Speech and language problems	0	0

Bolded items represent disruptive behavior treatment targets. Treatment target utilized at least once throughout the episode = Number of participants (and percent of sample) that had each treatment target endorsed at least once throughout the entire episode (i.e., yes or no). Treatment target utilization overall = Percent of Monthly Treatment Progress Summary (MTPS) reporting months within a youth's entire treatment episode that included each treatment target, averaged (with standard deviation) across the entire dataset

\*Treatment targets commonly written-in by therapists

that were considered PDEB for disruptive behavior youth 13 years or older across all settings (bolded in Table 4<sup>3</sup>). The PDEB predictor score was defined as the average number of PDEBs endorsed as being used by a therapist with a client per month (henceforth known as PDEB-score). For example, if a therapist saw a client in the CBR setting for 4 months and used 2 PDEBs in month one, 3 PDEBs in month two, 4 PDEBs in month three, and 5 PDEBs in month four, that youths PDEB-score would be  $(2 + 3 + 4 + 5)/4 = 3.5$ . PMES-score (i.e., every practice element other than the 15 PDEBs noted above) was defined as the average number of PMESs used per month and were calculated similarly to the PDEB-score. Each youth received one PDEB-score and one PMES-score.

### Data Preparation, Missing Data, and Power

Response ranges, means, standard deviations, skewness, and kurtosis were investigated for all measures. Since multilevel modeling assumes that the missing data in the sample are missing at random (Heck et al. 2013), missing data was analyzed and addressed as needed (see Results below). Regarding power, attempts were made to increase units at level-two (i.e., clients; Heck et al. 2013) and prior studies

using multilevel modeling in this field were referenced for appropriate sample sizes (e.g., between 57 to 2,171; Mueller et al. 2010; Orimoto et al. 2013; Weersing and Weisz 2002). Based on these factors, the final sample of 341 likely had sufficient power.

### Analyses for Study Aim One: Descriptive Analyses

In addition to the means, standard deviations, and frequencies that were calculated for youth, therapist, and clinical information displayed in Tables 1 and 2, the frequency counts of MTPS treatment targets used were calculated in two ways: (a) the proportion of the total sample that had each treatment target endorsed at any point in time throughout the entire treatment episode (regardless of how often; e.g., 1 = yes, 0 = no) and (b) the monthly proportion of all MTPSs in a youth's entire treatment episode that included each treatment target (e.g., 8 of 10 MTPSs = 80%), averaged across the entire dataset. The frequency counts for each practice element were calculated in the same fashion.

### Analyses for Study Aim Two: Multilevel Modeling

The multilevel modeling in the current study examined the extent to which the PDEB-score and PMES-score for disruptive youth in the CBR setting predicted rate of change or slope of youth progress ratings for their disruptive behavior problem treatment targets (i.e., the dependent variable). The final progress rating was calculated for each youth by averaging their MTPS progress ratings on the seven disruptive behavior targets for each month and then looking at the change between the first and last months. Treatment episode length for the multilevel modeling analyses was limited to the first 9 months of treatment (i.e., limited by the average

<sup>3</sup> As is evident from the PDEBs bolded in Table 4, a few practice elements target parents and teachers (i.e., parent and teacher monitor and praise). These PDEBs were kept in the list of PDEBs used to create the PDEB-score because families are typically incorporated into therapy sessions toward the end of a youth's stay to help with their transition back home. In addition, youth in residential settings still have teachers who are part of their treatment team within their agency. Hence, both parents and teachers play active roles in the team, but to a different extent as they would in other settings.

**Table 4** Frequency counts of practice elements for the total sample (n = 341)

Practice element (PE)	PE utilized at least once throughout episode	PE utilization overall
Supportive listening or client centered	335 (98.2%)	87.2% (20.5)
<b>Communication skills</b>	<b>340 (99.7%)</b>	<b>86.2% (19.5)</b>
Natural and logical consequences	329 (96.5%)	83.8% (24.7)
Line of sight supervision	322 (94.4%)	83.3% (28.4)
Time out	331 (97.1%)	83.0% (24.8)
Educational supports	329 (96.5%)	82.6% (26.3)
<b>Relationship or rapport building</b>	<b>337 (98.8%)</b>	<b>82.6% (21.7)</b>
Activity scheduling	329 (96.5%)	81.9% (28.6)
<b>Problem solving</b>	<b>335 (98.2%)</b>	<b>80.2% (24.2)</b>
Skill building	334 (97.9%)	79.2% (25.1)
Milieu therapy	310 (90.9%)	78.1% (32.7)
Insight building	329 (96.5%)	76.4% (27.8)
<b>Family therapy</b>	<b>319 (93.5%)</b>	<b>75.8% (29.5)</b>
Family engagement	328 (96.2%)	75.8% (28.9)
<b>Therapist praise or rewards</b>	<b>299 (87.7%)</b>	<b>73.6% (36.3)</b>
<b>Social skills training</b>	<b>312 (91.5%)</b>	<b>72.6% (32.4)</b>
<b>Maintenance or relapse prevention</b>	<b>307 (90.0%)</b>	<b>71.4% (33.4)</b>
Emotional processing	297 (87.1%)	69.1% (36.1)
Twelve step program	282 (82.7%)	68.6% (39.8)
<b>Cognitive</b>	<b>298 (87.4%)</b>	<b>68.6% (36.3)</b>
<b>Modeling</b>	<b>300 (88.0%)</b>	<b>63.8% (35.3)</b>
Crisis management	298 (87.4%)	63.2% (35.3)
<b>Tangible rewards</b>	<b>286 (83.9%)</b>	<b>62.7% (40.0)</b>
Psychoeducation child	301 (88.3%)	61.4% (36.6)
Mindfulness	291 (85.3%)	59.2% (36.8)
<b>Goal setting</b>	<b>243 (71.3%)</b>	<b>56.3% (41.8)</b>
<b>Self-monitoring</b>	<b>293 (85.9%)</b>	<b>53.5% (33.9)</b>
Care coordination	229 (67.2%)	53.1% (42.5)
Mentoring	264 (77.4%)	52.9% (38.7)
Assertiveness training	265 (77.7%)	52.6% (38.6)
Relaxation	268 (78.6%)	52.4% (38.0)
Self-rewards or self-praise	247 (72.4%)	49.1% (39.6)
Personal safety skills	217 (63.6%)	48.5% (42.9)
Motivational interviewing	246 (72.1%)	47.9% (40.8)
Commands	217 (63.6%)	46.1% (43.1)
Medication or pharmacotherapy	228 (66.9%)	45.3% (38.9)
Peer pairing	241 (70.7%)	41.6% (37.1)
<b>Parent or teacher praise</b>	<b>226 (66.3%)</b>	<b>37.8% (36.4)</b>
Psychoeducation parent	250 (73.3%)	37.7% (33.7)
Behavioral contracting	197 (57.8%)	35.9% (38.0)
Cultural training	170 (49.9%)	34.4% (42.0)
Ignoring or DRO	196 (57.5%)	34.3% (36.8)
Parent coping	198 (58.1%)	31.3% (34.3)
Attending	169 (49.6%)	31.2% (38.9)
<b>Parent or teacher monitoring</b>	<b>189 (55.4%)</b>	<b>29.7% (35.0)</b>
Interpretation	184 (54.0%)	26.4% (33.1)
Stimulus control/antecedent management	136 (39.9%)	21.5% (33.0)
Response prevention	124 (36.4%)	15.1% (26.6)
<b>Response cost</b>	<b>102 (29.9%)</b>	<b>13.5% (25.9)</b>
Free association	90 (26.4%)	12.3% (26.6)

**Table 4** (continued)

Practice element (PE)	PE utilized at least once throughout episode	PE utilization overall
Functional analysis	84 (24.6%)	11.1% (24.9)
Exposure	101 (29.6%)	11.1% (22.1)
Unclear*	55 (16.1%)	8.4% (23.4)
Guided imagery	35 (10.3%)	6.2% (21.6)
Play therapy	66 (19.4%)	5.6% (15.1)
Catharsis	51 (15.0%)	5.4% (16.8)
Other*	35 (10.3%)	5.0% (17.2)
Individual therapy for caregiver	52 (15.2%)	3.6% (11.7)
Anger management*	19 (5.6%)	2.2% (10.2)
Thought field therapy	27 (7.9%)	1.2% (4.4)
Biofeedback or neurofeedback	12 (3.5%)	0.8% (4.5)
Marital therapy	14 (4.1%)	0.5% (2.8)
Eye movement or tapping	6 (1.8%)	0.4% (4.7)
Discrete trial training	5 (1.5%)	0.4% (3.1)
Art or music therapy*	3 (0.9%)	0.3% (3.9)
Parenting*	1 (0.3%)	0.2% (2.7)
Counseling*	2 (0.6%)	0.1% (1.3)
Hypnosis	2 (0.6%)	0.1% (1.3)
Juvenile sex offender treatment*	1 (0.3%)	0.0% (0.5)

Bolded items indicate practice elements that derived from the evidence-base (i.e., PDEBs; defined earlier as practice elements in 30% or more of the study groups examining evidence-based treatments for youth 13 years or older across all settings). PE utilized at least once throughout the episode=Number of participants (and percent of sample) that had each practice element endorsed at least once throughout the entire episode (i.e., yes or no). PE utilization overall=Percent of Monthly Treatment Progress Summary reporting months within a youth's entire treatment episode that included each practice element, averaged (with standard deviation) across the entire dataset. DRO=Differential Reinforcement of Other Behaviors

\*Practice elements commonly written-in by therapists

length of treatment for a CBR episode plus one standard deviation). Statistical Package for Social Sciences was utilized to analyze the three-level<sup>4</sup> mixed-effects model, where time (in months) that the MTPS was completed was nested within youth, which was nested within therapists. Level-one included time in months. Level-two included the main variables of interest (i.e., PDEB- and PMES-scores) and controlled for between-client variation and youth-related variables. Level-three included therapist-level characteristics as covariates.

<sup>4</sup> This model was originally going to include four-levels, with the fourth level being CBR-III agency. However, due to the small number of agencies in this sample (n=5), it was removed as a level and was instead included as a variable at the client level (i.e., level-two).

## Results

### Data Preparation, Assumptions, and Missing Values

Ranges, means, standard deviations, and kurtosis were examined and found normal. Only nine MTPSs (0.4% of the total 2419 MTPSs) did not have practice elements and 33 MTPSs (1.4% of the total 2419 MTPSs) were missing both treatment target and practice elements. Due to multilevel modeling allowing for missing data within clients, no actions were taken. Multiple imputation was used to calculate missing values for the Child and Adolescent Functional Assessment Scale Total Score (i.e., 13.5% of the total sample) using relevant client-level variables. No other variables had missing data that was a concern. The following assumptions for multilevel modeling were tested and met when the sample was limited to the first nine months of the treatment episode (i.e., 2241 rows of data): the distribution of the dependent variable was normal; the residuals of the dependent variable beyond  $\pm$  three (55 data points) were removed; and multicollinearity was not a concern.

## Analyses for Study Aim One: Descriptive Analyses

The frequency counts of all treatment targets endorsed on the MTPSs throughout the entire treatment episode (i.e., prior to limiting the sample to nine months) for the entire sample are presented in Table 3 and interpreted as noted in the “Methods” section. For example, oppositional/non-compliant behavior was used at least once throughout the treatment episode for 286 youth (83.9% of the total sample) and was used on average within 76.0% ( $SD = 37.6$ ) of the months throughout all youths’ treatment episodes. Overall, the results from Table 3 suggested that the top disruptive behavior treatment targets across the total sample were anger, aggression, oppositional/non-compliant behavior, runaway, and willful misconduct/delinquency. In addition, these results indicated that fire setting and sexual misconduct were either of minimal or no concern for this sample. Based on the results from the total sample, non-disruptive behavior treatment targets that frequently occurred included the following: other externalizing behaviors (e.g., substance use), social interactions (e.g., activity involvement), anxiety (e.g., avoidance), mood (e.g., contentment/enjoyment/happiness), and other topics (e.g., treatment engagement).

The frequency count of practice elements endorsed on the MTPSs throughout the youth’s entire treatment episode (i.e., prior to limiting the sample to nine months) for the total sample are presented in Table 4. The columns in this table can be interpreted in the same fashion as the columns in Table 3. For example, in Table 4, supportive listening or client centered was used at least once throughout the treatment episode for 335 youth (98.2% of the sample) and was used on average within 87.2% ( $SD = 20.5$ ) of the months throughout all youths’ treatment episodes. Table 4 suggests that all of the PDEBs, except response cost, were used at least once throughout the entire treatment episode for over half of the total sample (i.e., communications skills, relationship or rapport building, problem solving, family therapy, therapist praise or rewards, social skills training, maintenance or relapse prevention, cognitive, modeling, tangible rewards, goal setting, self-monitoring, parent or teacher praise, and parent or teacher monitoring). In addition, all of the PDEBs, except parent or teacher monitoring, parent or teacher praise, and response cost, were used in at least 50% of the MTPS reporting months for the total sample. Based on this information, it appears that the most frequently occurring PDEBs primarily target skills implemented for older youth by therapists in individual therapy settings. However, after visually inspecting the data from Table 4, it was evident that the overall majority of the practice elements that were utilized in over 50% of the months throughout all youths’ treatment episodes were targeted to children (e.g., activity scheduling, assertiveness training, cognitive, communication skills, emotional processing).

## Analyses for Study Aim Two: Multilevel Modeling

For the multilevel modeling analyses, full information maximum likelihood was chosen as a parameter estimation method so that successive models with both regression coefficients and variance components could be compared. Due to this multilevel modeling being longitudinal, the ICC was calculated with the level-one model that included time (Heck et al. 2013). Level-one, -two, and -three accounted for 30.99%, 34.50%, and 34.50% of the variance in this initial model, suggesting that there was sufficient variance (more than 5%) at level-two (i.e., the grouping variable) to conduct multilevel modeling. When inspecting the within-subjects growth trends for a random sample of 30 subjects, the majority of the trends appeared to be linear, suggesting that it was appropriate to keep time as linear. The intercept was also recoded<sup>5</sup> as ending status (i.e., the level of the dependent variable at the end of treatment, adjusted for covariates in the model) so that the intercept was interpreted as the final average progress or improvement rating on MTPS disruptive behavior treatment targets and time was interpreted as the change from the first to last month of treatment. Finally, to help with interpretation of the data, several variables were centered on their sample average (i.e., grand-mean centering) or minimum value.<sup>6</sup>

### Intercept-Only Model

The intercept-only model (no time included) indicated that the average progress rating across months was significantly different than zero ( $p < 0.001$ ). The intercept of 2.63 (i.e., “minimal” to “some improvement”) was the grand mean progress rating on disruptive behavior treatment targets across all months and all clients. A Scaled Identity error structure fit the level-one data best (Heck et al. 2013).

### Time-Only (Level-One) Model

This model included the time variable added as a fixed effect at level-one and random effects for time at level-two and -three. The end status intercept of this model was 3.83

<sup>5</sup> To define the intercept as ending status, the time variable was coded such that the last month of treatment was 0, and the first month of treatment was -1. The months of treatment between -1 and 0 varied depending on the length of youth’s treatment episode. For example, a client with 4 months was coded -1, -0.66, -0.33, and 0, while a client with 9 months was coded -1, -0.825, -0.75, -0.675, -0.5, -0.325, -0.25, -0.175, and 0.

<sup>6</sup> Variables were either centered around the grand mean or their minimum values in order to facilitate interpretation of the intercept. These decisions to center variables did not change their slopes; hence, the choices were made based on how best to aide in the interpretation of the intercept.

**Table 5** Multilevel models predicting end status and monthly rate of change in average progress rating on disruptive behavior treatment targets using average number of PDEBs and PMESs (n = 341)

		Level-one model	Level-two model
<b>Fixed effects</b>			
Final average progress rating	Intercept	3.83** (SE=0.14)	2.44** (SE=0.31)
	Length of treatment beyond 3 months		0.19** (SE=0.02)
	Average Sum of PDEB (CM)		0.08* (SE=0.04)
	Average Sum of PMES (CM)		0.03~ (SE=0.02)
	Agency A versus Agency B		-0.13 (SE=0.33)
	Agency C versus Agency B		-0.85 (SE=0.56)
	Agency D versus Agency B		0.39 (SE=0.40)
	Agency E versus Agency B		-1.06* (SE=0.50)
	Age in months (GMC)		-0.01** (SE=0.00)
	Time	2.46** (SE=0.12)	2.43** (SE=0.13)
<b>Variance components</b>			
Level-1 (time)	Within-person	0.53**	0.51**
Level-2 (client)	In final status	0.59**	0.45**
	In rate of change	0.19~	0.28*
Level-3 (therapist)	In final status	0.59**	0.51**
	In rate of change	0.50**	0.53**
<b>Goodness of fit</b>			
	Deviance	5726.90	5624.77
	No of estimated parameters	7	15
	AIC	5740.90	5654.77
	BIC	5780.73	5740.11

*PDEB* Practices derived from the evidence-base, *PMES* practices with minimal evidence support, *GMC* grand-mean centered, *CM* centered on the minimum, *SE* standard error, *AIC* Akaike information criterion, *BIC* Bayesian information criterion

~ $p < 0.10$ ; \* $p < 0.05$ ; \*\* $p < 0.001$

( $p < 0.001$ ), suggesting that at the final month of their treatment episode, participants' final average progress rating on disruptive behavior treatment targets was 3.83 on the 0 to 6 point scale (i.e., "some" to "moderate improvement") when only time is considered as a predictor. The estimate for time was 2.46 ( $p < 0.001$ ), suggesting that across the varied lengths of their treatment episodes, the average change in the mean progress rating on disruptive behavior treatment targets increased by 2.46 on the MTPS rating scale. The covariance parameters of this model were examined within youth (Wald  $Z = 18.82$ ,  $p < 0.001$ ), between youth (Wald  $Z = 6.27$ ,  $p < 0.001$ ), and between therapists (Wald  $Z = 3.093$ ,  $p < 0.002$ ). Random effects were 0.18 ( $p = 0.091$ ) at level-two and 0.50 ( $p = 0.002$ ) at level-three. The  $-2$  Log Likelihood deviance value for the final level-one model was 5726.90, suggesting this model was considerably better than the intercept-only model ( $-2$  Log Likelihood = 7425.64).

### Level-Two Model

In addition to the time variable from the level-one model, the following variables were added into the model as fixed effects for explaining the level-two ending status intercept

(i.e., final mean progress rating on disruptive behavior targets): PDEB-score, PMES-score, length of treatment beyond three months<sup>7</sup>, agency (with the comparison group being the agency with the largest sample size: Agency B), gender, age in months (centered on the grand mean), total Child and Adolescent Functional Assessment Scale score at the start of the treatment episode (centered on the grand mean), and number of diagnoses. After removing the insignificant predictors and only keeping variables significant at  $p < 0.10$  for exploratory purposes, the final level-two model (see Table 5) for the end status intercept included the significant variables of PDEB-score (centered on the minimum), length of treatment beyond three months, age in months (centered on the grand mean), and agency (with the comparison group being Agency B). In addition, the variable that was significant

<sup>7</sup> It should be noted that the sample was initially restricted to those youth who had two or more months of CBR treatment to ensure that there was enough data points to see a change in progress rating over time. However, after the remainder of the inclusionary criteria were applied (see "Participants" section), the final sample inadvertently had a minimum of 3 months of treatment.

at  $p < 0.10$  in this model included the PMES-score (centered on the minimum). These variables together changed the end status intercept to be 2.44 ( $p < 0.001$ ), which meant that the final average progress rating on disruptive behavior treatment targets was 2.44 (i.e., “minimal” or “some improvement” from baseline) for youth who had 3 months of treatment, had the minimum average number of PDEBs and PMESs used per month, were in Agency B (i.e., the comparison group), and had the average age in the sample.

More specifically, with every added month of treatment beyond three months, the final average progress rating on disruptive behavior treatment targets increased by 0.19 points on the MTPS ( $p < 0.001$ ). For every one-point increase in average number of PDEBs used per month beyond the minimum of 3.89, the final average progress rating on disruptive behavior treatment targets increased by 0.08 points ( $p = 0.036$ ). Finally, for every added month of age beyond the grand mean (i.e., 189.6 months or 15.8 years), the final average progress rating on disruptive behavior treatment targets decreased by 0.01 ( $p < 0.001$ ). Among the agencies included in this study, only Agency E was significant at  $p = 0.044$  in differing from Agency B. Youth in Agency E (which served both males and females) had 1.06 points less than youth in Agency B (which served only males) on their final average progress rating on disruptive behavior treatment targets ( $p = 0.044$ ). With every one-unit increase in average number of PMESs used per month beyond the minimum of 7.29, the final average progress rating on disruptive behavior treatment targets increased by 0.03 ( $p = 0.075$ ). Random effects of time were 0.44 ( $p = 0.002$ ) at level-three and 0.06 ( $p = 0.333$ ) at level-two. The deviance value for the final level-two model ( $-2 \text{ Log Likelihood} = 5624.77$ ) was significantly smaller than the previous model ( $-2 \text{ Log Likelihood} = 5726.90$ , over the Chi square critical value of 102.13 for  $df = 8$ ) suggesting that this level-two model was better at predicting the end status of average progress rating for disruptive behavior treatment targets.

### Level-Three Model

In addition to carrying-over the time variable from the level-one model and the significant level-two predictors, the following variables were added into the model as fixed effects at the level-three intercept model: highest educational degree, number of degrees, professional specialty, licensure status (i.e., yes or no), and position (i.e., Qualified Mental Health Professional, Mental Health Professional, and Paraprofessional). Given that none of the level-three variables were found to be a significant predictor of the ending status intercept, they were not included in the final model, resulting in the level-two model being selected as the final model.

## Discussion

This is the first study to date to comprehensively describe treatment target and practice characteristics of youth in CBR settings, as well as the extent to which therapeutic practices impact the improvement rates of older youth with disruptive behavior problems within this setting. The study’s first aim was to conduct exploratory descriptive analyses to better understand the practices used in CBR settings. Based on the results from the practice element utilization frequency counts, both PDEBs and PMESs were used with CBR youth. Said another way, therapists in CBR settings reported using a wide variety of practices, those both associated and not typically associated with the evidence-base for disruptive behaviors. The most commonly used practice elements across the entire sample (i.e., 90% or more) were typically child-oriented. The specific PDEBs that were consistently not administered among the majority of the youth in the sample were parent or teacher monitoring, parent or teacher praise, and response cost. Regarding treatment targets, the non-disruptive behavior treatment targets covered a variety of issues including social interactions, anxiety, mood, and other topics.

The study’s second aim was to determine the extent to which PDEB-use and PMES-use predicted improvement rates for older youth with disruptive behaviors in the CBR setting. The hypothesis for the second aim of this study was supported: both PDEB-score and PMES-score were significant at  $p < 0.05$  and  $p < 0.10$ , respectively, in positively predicting final average progress rating. Along with these predictors, lower age, longer length of treatment, and agency were significant in predicting final average progress rating for disruptive behavior targets. In other words, youth improved at better rates when more practices were used with them (i.e., practices that both aligned and did not align with the evidence-base), they were younger, they have longer treatment episodes, and they were part of certain agencies.

Taken as a whole, the findings in which PDEB-score predicted final average progress rating on disruptive behavior treatment targets align with previous research studies which found PDEBs (whether measured by proportionate value or sum of frequencies) was significant at predicting either final average progress rating or improvement rates on progress ratings for depressed, disruptive, and inattention/hyperactive youth in various settings (Love 2014; Mueller et al. 2009; Orimoto et al. 2013). However, it is important to interpret this aforementioned finding with its complementary result, which noted that the PMES-score was marginally significant (i.e.,  $p < 0.10$ ) in positively predicting final average progress rating on disruptive behavior treatment targets. This

might suggest that therapists are using practice elements that derive from the evidence-base for multiple problem areas (e.g., exposure for anxiety, problem solving for disruptive behavior), which might be the reason both the PDEB-score and PMES-score were significant and marginally significant, respectively, in positively predicting final average progress rating. As can be seen from the demographic information and the treatment target frequencies, most youth in this sample had multiple problem areas beyond that of disruptive behavior. Hence, it may be that these therapists were mindful of the comorbidity of their clients when picking PDEBs for disruptive behavior and other problem areas (although PDEBs for other problem areas would be classified as disruptive behavior PMES strategies). This interpretation also relates to previous research that found that youth with two or more diagnoses had an eclectic range and greater dosage of practice elements throughout their treatment (Orimoto et al. 2014). Another interpretation of these results is that the therapists are using a variety of practices with their clients, without choosing them based on the evidence-based literature (i.e., they are “throwing the kitchen sink” at them). These therapists may be unsure of how to treat these severe clients and may turn to their previous work with youth in these settings to guide the practices they used. In addition, it may be likely that the therapists are using a more flexible approach to using evidence-based practices, instead of only following brand named treatment protocols. Overall, it appears that these results as a whole can be interpreted as either therapists consciously using PDEBs from multiple problem areas, therapists using as many practice elements as they deem fit for their clients, or therapists referencing their prior experiences with these clients to guide their treatment.

### Client- and Therapist-Level Variables Predicting Improvement

Alongside the influence of practice elements, lower youth age and longer length of treatment were significant in predicting final average progress rating on disruptive behavior treatment targets. In other words, younger youth who stayed in the residential setting had better outcomes in treatment. Lower age was potentially related to better progress ratings since some of the older youth might have “aged out” of the mental health system during their CBR treatment episode and might have chosen to leave the program regardless of whether or not they improved. Another interpretation of this result is that these older youth may have had less PDEBs used with them, leading them to not improve at the same rate as younger clients (who in the Child and Adolescent Mental Health Division system, have been shown to receive less PMESs than older youth; Higa-McMillan et al. 2014). Regarding longer length of treatment, the simplest interpretation of this finding is that more practice elements or

PDEBs can be used with clients and mastered if they have longer lengths of treatment, which increases the likelihood of them improving. To take this interpretation a step forward, some of these CBR agencies have treatment programs or a specific set of therapeutic topics that they are required to cover with each youth (e.g., substance abuse counseling). Thus, if a youth had a longer length of treatment, they likely progressed through more of the agency’s materials, leading them to learn more therapeutic skills and improve at higher rates. Additionally, youth with shorter episodes might have quit the program (e.g., aging out, eloping), leaving them without as many therapeutic skills. However, these results should be interpreted with caution due to the sample being restricted to those with three or more months of treatment after all of the inclusionary criteria were applied. This sample restriction might have had inadvertent effects on the results that make it difficult to discuss the youth with 1–2 months of treatment in a CBR setting. On a different note, anecdotal information provided by one agency suggested that some youth are encouraged to stay in their program long after they have met their treatment goals due to their families and homes not being deemed as appropriate for their return. Thus, it may be that youth who have longer lengths of treatment have maintained the improvement from prior months and are just practicing their skills in a controlled, CBR environment. Finally, results related to agency suggest that an agency’s specialty population might also impact the improvement rate of youth, with agencies that service broader youth (i.e., both males and females) improving at lower rates than agencies that serve smaller subsets of youth. This may be due to agencies servicing narrower populations being able to identify specific practice elements that are more effective with their youth.

Regarding therapist data, the current study found that no therapist-specific information was significant in predicting final average progress rating. Previous research in Hawai‘i that investigated these same constructs (e.g., specialty, position, degree, licensure status) found similar non-significant results (Love 2014; Orimoto et al. 2013). These findings might have occurred for several reasons. First, the categories of the specialties included in this study were derived based on the highest degree held by the therapist. Nuanced information collected directly from therapists might have led to different categorizations of professional specialties, which may have been significant in predicting final status or rate of change. Second, this data may not have had enough variance at the therapist level to indicate significant therapist-level predictors. For example, since the majority of the therapists had a masters as their highest degree, it decreased the chances of it being significant in predicting the outcome. A third reason might relate to the lack of important therapist-related constructs not examined in this study that were found significant in previous research (e.g., theoretical orientation,

years of experience, evidence-based practice attitudes; Brookman-Frazee et al. 2010). Finally, these insignificant therapist variables might suggest a deeper examination of service setting or organizational characteristics (e.g., leadership, culture, mission, resources) as a predictor of improvement, given the recent shift toward examining these factors in the evidence-based practice implementation movement (e.g., Aarons et al. 2011) and agency differences suggested by these results. Taken together, having additional therapist data that was more refined could have led to in depth analyses that might have significantly predicted youth progress rating over time.

### Limitations

Although the results of the present study are promising with regard to the influence of PDEB-use, lower age, and length of treatment on the final average progress rating on disruptive behavior treatment targets, limitations should be considered. First, several variables from this study were based on data collected from the MTPS, a self-report measure filled out by the therapist, which are at risk of having a reporter-bias, especially when they might be perceived as reflecting the quality of work completed by the reporter. Second, there was a limited number of therapist-and client-related variables included in this study. This is especially true given the several therapists that may have worked with these youth in the CBR setting. Third, CBR agencies may specialize in specific populations, which might make the generalizability of these findings dependent on the specific agency. Further expanding upon this point, given Hawai'i's extensive and long history regarding youth mental health quality improvement, some caution may be warranted with regard to generalizing these findings to other mental health systems. Fourth, due to attempts to maximize sample size, youth with various lengths of treatment episodes and discharge statuses were included, which may have impacted the results. In addition, care and caution should be used in interpreting the study's results with respect to specificity of the observed effects regarding both PDEBs and PMESs being positively related to the improvement rate for youth with disruptive behaviors due to the high rate of comorbidity and the complex clinical presentations of these youth. Results might have differed if PMES and PDEB metrics were based on each youth's unique comorbidity presentation. Finally, the method of determining PDEBs (i.e., derived from protocols for older youth with disruptive behaviors across all settings) may have been too broad of a PDEB profile for CBR youth.

### Implications and Future Studies

There are several implications for these results. Overall, the innovative methodology accompanied with the specific

patterns of findings in this paper highlight the potential necessity of continuing to acknowledge the importance of avoiding an all-or-nothing approach when it comes to large-scale quality improvement and evidence-based practice implementation (Nakamura et al. 2014). It is important to consider implementing evidence-based practices in a variety of ways, especially when it relates to comorbid youth and youth that might be placed in unique settings that are not heavily researched (e.g., CBR). Therapists should consider the importance of being "guided" by the research when implementing evidence-based practices, as opposed to being limited by specific manuals or trainings that only focus on a specific diagnosis or problem area. Future studies may elaborate on these findings by further investigating youth comorbidity, how this relates to the use of practice elements (e.g., PDEB- and PMES-use) and the improvement rates for youth in this setting, and whether comorbidity is truly the reason for the high use of PMESs among therapists working with disruptive youth in CBR settings. More qualitative information regarding the clinical decision-making process of therapists (e.g., barriers, agency-requirements about treatment programs) may also provide information about PDEB- and PMES-use. Future research should continue to investigate treatment as usual practices of the severe youth placed in the CBR setting (e.g., examining specific PDEBs derived from CBR protocols, non-disruptive behavior diagnoses, youths' insight into practice elements that were helpful) or add to the limited number of randomized controlled trials in this setting. In addition, investigating factors related to maintaining improvement post-discharge from a CBR setting would also be beneficial to the field. Finally, future studies should continue to investigate treatment as usual using multiple methods of defining evidence-based practices or PDEBs (e.g., using different percentage criteria for defining PDEBs, average number used per month, proportionate use, treatment manuals).

Despite these limitations and indications for future research, the present study was the first investigation of the extent to which PDEB-use predicted improvement for youth with disruptive behaviors in CBR settings. Overall, the findings suggested that the average number of PDEBs used per month, average number of PMESs used per month, lower age of clients, longer length of treatment, and agency were significant or marginally significant in predicting final average progress rating. Given the importance of honing in on the specific practices that help youth from CBR settings improve, advancing empirical inquiry into treatment as usual practices for these highly impaired population continues to be a worthwhile endeavor.

**Funding** This study did not receive any funding.

## Compliance with Ethical Standards

**Conflict of interest** Dr. Izmirian and Dr. Chang declare that they have no conflict of interests. Dr. Nakamura has received funding from the State of Hawaii Child and Adolescent Mental Health Division, and does a small amount of consulting for PracticeWise, LLC.

**Ethical Approval** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

**Informed Consent** Informed consent was obtained from all individual participants included in the study.

**Research Involved in Human and Animal Rights** This article does not contain any studies with animals performed by any of the authors.

## 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 and Mental Health Related Services*, 38, 4–23. <https://doi.org/10.1007/s10488-010-0327-7>.
- Baumann, B. L., Kolko, D. J., Collins, K., & Herschell, A. D. (2006). Understanding practitioners' characteristics and perspectives prior to the dissemination of an evidence-based intervention. *Child Abuse & Neglect*, 30(7), 771–787. <https://doi.org/10.1016/j.chiabu.2006.01.002>.
- Bickman, L. (2000). The most dangerous and difficult question in mental health services research. *Mental Health Services Research*, 2(2), 71–72. <https://doi.org/10.1023/A:1010100119789>.
- 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. *Journal of Behavioral Health Services and Research*, 42(3), 367–382. <https://doi.org/10.1007/s11414-013-9363-x>.
- Brookman-Frazee, L., Haine, R. A., Baker-Ericzén, M., Zoffness, R., & Garland, A. F. (2010). Factors associated with use of evidence-based practice strategies in usual care youth psychotherapy. *Administration and Policy in Mental Health and Mental Health Services Research*, 37(3), 254–269. <https://doi.org/10.1007/s10488-009-0244-9>.
- Child and Adolescent Mental Health Division. (2005). *Service provider monthly treatment and progress summary*. Honolulu: Hawaii Department of Health, Child and Adolescent Mental Health Division.
- Child and Adolescent Mental Health Division, Hawaii Department of Health Child and Adolescent Mental Health Division. (2012). *Child and Adolescent Mental Health Performance Standards Interagency Performance Standards and Practice*. Retrieved from <http://health.hawaii.gov/camhd/>.
- Chorpita, B. F., & Daleiden, E. (2009). *Evidence-based services committee—Biennial report—Effective psychological interventions for youth with behavioral and emotional needs*. Honolulu: Hawaii Department of Health Child & Adolescent Mental Health Division.
- Chorpita, B. F., Daleiden, E. L., & Weisz, J. R. (2005). Identifying and selecting the common elements of evidence-based interventions: A distillation and matching model. *Mental Health Services Research*, 7, 5–20. <https://doi.org/10.1007/s11020-005-1962-6>.
- Daleiden, E. L., Lee, J., & Tolman, R. (2004). *Annual report*. Honolulu, HI: Hawaii Department of Health, Child and Adolescent Mental Health Division.
- Daleiden, E. L., Pang, D., Roberts, D., Slavin, L. A., & Pestle, S. L. (2010). Intensive home based services within a comprehensive system of care for youth. *Journal of Child and Family Studies*, 19, 318–325. <https://doi.org/10.1007/s10826-009-9300-z>.
- Denenny, D., & Mueller, C. (2012). *Do empirically supported packages or their practices predict superior therapy outcomes for youth with conduct disorders?* Paper presented at the forty-sixth annual convention of the association of behavioral and cognitive therapies, National Harbor, MD.
- 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. <https://doi.org/10.1176/appi.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 and Adolescent Psychiatry*, 47(5), 505–514. <https://doi.org/10.1097/CHI.0b013e31816765c2>.
- Garland, A. F., Hurlburt, M. S., & Hawley, K. M. (2006). Examining psychotherapy processes in a services research context. *Clinical Psychology: Science and Practice*, 13, 30–46. <https://doi.org/10.1111/j.1468-2850.2006.00004.x>.
- Heck, R. H., Thomas, S. L., & Tabata, L. N. (2013). *Multilevel and longitudinal modeling with IBM SPSS: Quantitative methodology series*. (second edition). New York: Routledge.
- Higa-McMillan, C., Powell, C. K. K., Daleiden, E. L., & Mueller, C. W. (2011). Pursuing an evidence-based culture through contextualized feedback: Aligning youth outcomes and practices. *Professional Psychology: Research & Practice*, 42(2), 137–144. <https://doi.org/10.1037/a0022139>.
- Higa-McMillan, C. K., Nakamura, B. J., Morris, A., Jackson, D. S., & Slavin, L. (2014). Predictors of use of evidence-based practices for children and adolescents in usual care. *Administration and Policy in Mental Health and Mental Health Services Research*, 42(4), 373–383. <https://doi.org/10.1007/s10488-014-0578-9>.
- Hoagwood, K., & Kolko, D. J. (2009). Introduction to the special section on practice contexts: A glimpse into the nether world of public mental health services for children and families. *Administration and Policy in Mental Health and Mental Health Services*, 36(1), 35–36. <https://doi.org/10.1007/s10488-008-0201-z>.
- Hodges, K. (1990, 1994 revision). *Child and adolescent functional assessment scale*. Ypsilanti: Eastern Michigan University, Department of Psychology.
- Hodges, K., & Gust, J. (1995). Measures of impairment for children and adolescents. *Journal of Mental Health Administration*, 22(4), 403–413. <https://doi.org/10.1007/BF02518634>.
- Hodges, K., & Wong, M. M. (1996). Psychometric characteristics of a multidimensional measure to assess impairment: The Child and Adolescent Functional Assessment Scale. *Journal of Child and Family Studies*, 5(4), 445–467. <https://doi.org/10.1007/BF02233865>.
- Izmirian, S. C., Nakamura, B. J., Hill, K. A., Higa-McMillan, C. K., & Slavin, L. (2016). Therapists' knowledge of practice elements derived from the evidence-base: Misconceptions, accuracies, and large-scale improvement guidance. *Professional Psychology: Research & Practice*. <https://doi.org/10.1037/pro0000081>. Advanced online publication.
- Kazdin, A. E. (1995). Child, parent and family dysfunction as predictors of outcome in cognitive-behavioral treatment of antisocial

- children. *Behaviour Research and Therapy*, 33, 271–281. [https://doi.org/10.1016/0005-7967\(94\)00053-M](https://doi.org/10.1016/0005-7967(94)00053-M).
- Keir, S. S., Jackson, D. S., Mueller, C. W., & Wilkie, D. (2014). *Child and adolescent mental health division: Fiscal year 2013 annual factbook*. Retrieved from Hawaii Department of Health Child and Adolescent Mental Health Division website: <http://health.hawaii.gov/camhd/resource-library/research-and-evaluation/>.
- Love, A., Orimoto, T., Powell, A., & Mueller, C. (2011, November). *Examining the relationship between therapist-identified treatment targets and youth diagnoses using exploratory factor analysis*. Paper presented at the Forty-Fifth Annual Convention of the Association of Behavioral and Cognitive Therapies, Toronto, Ontario.
- Love, A. R. (2014). *A Multilevel longitudinal analysis of the relationship between therapist use of practices derived from the evidence base and outcomes for youth with mood disturbances in community-based mental health* (Unpublished doctoral dissertation). University of Hawai'i at Mānoa, Honolulu, Hawaii.
- Lyons, J. S., & McCulloch, J. R. (2006). Monitoring and managing outcomes in residential treatment: Practice-based evidence in search of evidence-based practice. *Journal of the American Academy of Child and Adolescent Psychiatry*, 45(2), 247–251. <https://doi.org/10.1097/01.chi.0000190468.78200.4e>.
- Mueller, C., Daleiden, E., Chorpita, B., Tolman, R., & Higa-McMillan, C. (2009, August). Evidence-based service practice elements and youth outcomes in a state-wide system. In A. Marder (Chair), *A demonstration of mapping and traversing the science-practice gap*. Symposium conducted at the American Psychological Association, Toronto, Canada.
- Mueller, C. W., Tolman, R., Higa-McMillan, C. K., & Daleiden, E. L. (2010). Longitudinal predictors of youth functional improvement in a public mental health system. *The Journal of Behavioral Health Services & Research*, 37(3), 350–362. <https://doi.org/10.1007/s11414-009-9172-4>.
- 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. <https://doi.org/10.1007/s10826-006-9119-9>.
- Nakamura, B. J., Mueller, C. W., Higa-McMillan, C., Okamura, K. H., Chang, J. P., Slaving, L., & Shimabukuro, S. (2014). Engineering youth services system infrastructure: Hawaii's continued efforts at large-scale implementation through knowledge management strategies. *Journal of Clinical Child and Adolescent Psychology*, 43(2), 179–189. <https://doi.org/10.1080/15374416.2013.812039>.
- Norcross, J. C., & Karpiak, C. P. (2012). Clinical psychologists in the 2010s: 50 years of the APA division of clinical psychology. *Clinical Psychology: Science and Practice*, 19(1), 1–12. <https://doi.org/10.1111/j.1468-2850.2012.01269.x>.
- Orimoto, T. E., Higa-McMillan, C. K., Mueller, C. W., & Daleiden, E. L. (2012a). Assessment of therapy practices in community treatment for children and adolescents. *Psychiatric Services*, 63(4), 343–350. <https://doi.org/10.1176/appi.ps.201100129>.
- Orimoto, T. E., Jackson, D., Keir, S., Ku, J., & Mueller, C. W. (2012b). *Is less more? Patterns in client change across levels of care in youth public health services*. Paper presented at the Association for Behavioral and Cognitive Therapies, National Harbor, MD.
- Orimoto, T. E., Mueller, C. W., Hayashi, K., & Nakamura, B. J. (2014). Community-based treatment for youth with co- and multimorbid disruptive behavior disorders. *Administration and Policy in Mental Health and Mental Health Services Research*, 41(2), 262–275. <https://doi.org/10.1007/s10488-012-0464-2>.
- Orimoto, T. E., Mueller, C. W., & Nakamura, B. J. (2013). Fusing science and practice: Predicting progress ratings for disruptive behavior targets with practices derived from the evidence-based in community treatments. In R. Beidas & B. McLeod (Chair) and M. Southam-Gerow (Discussant), *Translating evidence-based assessment and treatments for youth for deployments in community settings*. Symposium presented at the Association for Behavioral and Cognitive Therapies, Nashville, TN.
- Pardini, D., & Fite, P. (2010). Symptoms of conduct disorder, oppositional defiant disorder, attention-deficit/hyperactivity disorder, and callous-unemotional traits as unique predictors of psychosocial maladjustment in boys: Advancing an evidence base for DSM-V. *Journal of the American Academy of Child & Adolescent Psychiatry*, 49, 1134–1144. <https://doi.org/10.1016/j.jaac.2010.07.010>.
- PracticeWise, L. L. C. (2014). PracticeWise evidence-base services database. Retrieved from <https://www.practicewise.com>.
- Soni, A. (2009). The five most costly children's conditions, 2006: Estimates for the U.S. civilian noninstitutionalized children, ages 0–17. Statistical Brief #242. Agency for healthcare research and quality, Rockville, MD, April 2009. [http://www.meps.ahrq.gov/mepsweb/data\\_files/publications/st242/stat242.pdf](http://www.meps.ahrq.gov/mepsweb/data_files/publications/st242/stat242.pdf).
- Southam-Gerow, M. A., Daleiden, E. L., Chorpita, B. F., Bae, C., Mitchell, C., Faye, M., & Alba, M. (2013). MAPping Los Angeles Country: Taking an evidenced-informed model of mental health care to scale. *Journal of Clinical Child & Adolescent Psychology*, 43(2), 190–200. <https://doi.org/10.1080/15374416.2013.833098>.
- Weersing, V. R., & Weisz, J. R. (2002). Community clinic treatment of depressed youth: Benchmarking usual care against CBT clinical trials. *Journal of Consulting and Clinical Psychology*, 70(2), 299–310. <https://doi.org/10.1037/0022-006X.70.2.299>.
- Weisz, J. R., Chorpita, B. F., Palinkas, L. A., Schoenwald, S. K., Miranda, J., Bearman, S. K., & ... The Research Network on Youth Mental Health (2012). Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth - A randomized effectiveness trial. *Archives of General Psychiatry*, 69(3), 274–282. <https://doi.org/10.1001/archgenpsychiatry.2011.147>.
- Weisz, J. R., & Jensen, A. L. (2001). Child and adolescent psychotherapy in research and practice contexts: Review of the evidence and suggestions for improving the field. *European Child and Adolescent Psychiatry*, 10(1), 112–118.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.