Child STEPs in California: A Cluster Randomized Effectiveness Trial Comparing Modular Treatment With Community Implemented Treatment for Youth With Anxiety, Depression, Conduct Problems, or Traumatic Stress

Bruce F. Chorpita
University of California, Los Angeles

Eric L. Daleiden
PracticeWise, LLC, Satellite Beach, Florida

Alayna L. Park
University of California, Los Angeles

Alyssa M. Ward
Hathaway-Sycamores Child and Family Services,
Los Angeles, California

Michelle C. Levy
University of California, Los Angeles

Taya Cromley
PracticeWise, LLC, Satellite Beach, Florida

Angela W. Chiu
Weill Cornell Medical College, Cornell University

Andrea M. Letamendi
Hathaway-Sycamores Child and Family Services,
Los Angeles, California

Katherine H. Tsai
Five Acres–The Boys' and Girls' Aid Society of
Los Angeles County, Los Angeles, California

Jennifer L. Krull
University of California, Los Angeles

Objective: This study reports outcomes from a randomized effectiveness trial testing modular treatment versus multiple community-implemented evidence-based treatments for youth. *Method:* An ethnoracially diverse sample of 138 youth ages 5 to 15 (62 girls, 76 boys) whose primary clinical concerns involved diagnoses or clinical elevations related to anxiety, depression, disruptive behavior, and/or traumatic stress were treated by community therapists randomly assigned to 1 of 2 conditions: (a) modular treatment, which involved a single modular protocol (i.e., modular approach to therapy for children; MATCH) that allowed flexible selection and sequencing of procedures to fit the chosen treatment focus in the context of measurement feedback, and (b) community-implemented treatment (CIT), which was a county-supported implementation of multiple evidence-based practices for youth. *Results:* Youth treated with MATCH showed significantly faster rates of improvement over time on clinical and functional outcomes

relative to youth in the CIT condition and required significantly fewer sessions delivered over significantly fewer days. Caregiver-reported clinical improvement rates were significantly greater for MATCH (60%) versus CIT (36.7%). Further, youth in the CIT condition were significantly more likely to receive additional psychosocial treatment services and were significantly more likely to use a variety of psychotropic medications during the active treatment phase. **Conclusions:** These results extend prior findings, supporting the effectiveness and efficiency of a modular, multifocus approach that incorporates monitoring and feedback relative to community implementation of evidence-based treatments.

What is the public health significance of this article?
This study shows that a modular, multifocus treatment program yields superior rates of improvement on clinical, functional, and utilization outcomes for youth receiving community mental health services. A coordinated and comprehensive treatment system, involving regular feedback and consultation shows advantages in terms of effectiveness and efficiency relative to standard implementation of multiple evidence-based treatments for youth.

Despite an impressive accumulation of research demonstrating the benefits of specific mental health treatments for youth in controlled research trials (e.g., Chorpita et al., 2011; Hoagwood et al., 2000; Nathan & Gorman, 2007; Weisz, Jensen-Doss, & Hawley, 2006), there continues to be a clear need to increase the impact of these treatments on those most in need (e.g., Kazdin & Blase, 2011; Miranda et al., 2005; Rotheram-Borus, Swendemann, & Chorpita, 2012). Research on dissemination and implementation of effective practices into community settings has helped to articulate some of the challenges (e.g., Aarons & Sawitzky, 2006; Proctor & Rosen, 2008) and possible solutions (e.g., Bickman, 2008; Glisson et al., 2008; Henggeler, Schoenwald, & Pickrel, 1995) to the problem of improving the lives of youth with mental health concerns.

One strategy in particular has focused on increasing the fit of intervention technologies with the service context through an innovative conceptualization of treatment design (Chorpita & Daleiden, 2014). Consistent with the notion that sustained use of an innovation depends heavily on its utility for the user (Taylor & Todd, 1995), this approach to treatment design structures the treatment protocol to be both developmental (e.g., meeting youths where they are, e.g., by giving them only those procedures that fit their treatment priorities) and dynamic (i.e., adapting in the face of emergent challenges, while remaining grounded in structured content and processes). This design approach ultimately elevates the role of therapist and family as collaborative designers of the treatment as it unfolds, along with the researcher/treatment developer, whose role is to provide the codified library of structured procedures and guidelines for their selection, use, and evaluation.

The notion that such protocol designs could perform well in actual service contexts was recently tested in a well-controlled experiment. In a randomized effectiveness trial conducted by the Research Network on Youth Mental Health (Weisz et al., 2012), therapists serving youth ages 7 to 13 with anxiety, depression, or conduct problems were assigned to one of three conditions: (a) modular treatment, which employed a flexible, components-based design that was guided by clinical algorithms and weekly feedback on practice and progress history; (b) standard evidence-based treatment (EBT), which involved training in three manualized evidence-based treatments; and (c) usual care (UC). Results from this initial trial showed that youth in the modular condition improved more quickly than youth receivingsStandard EBTs or UC on weekly measures of internalizing and externalizing symptoms and on weekly severity ratings of family concerns identified at the beginning of treatment. Youth in the modular condition also showed significantly fewer diagnoses at posttreatment than did youth in the UC condition, with standard EBTs not significantly differing from either of the other two conditions (Weisz et al., 2012). In addition, the treatment gains evidenced by the modular treatment were not only obtained more swiftly, but they also held up better over the long term, such that youth in the modular condition showed significantly greater improvement over two years relative to those in the UC condition (Chorpita et al., 2013).

Consistent with the aims of the effectiveness trial, therapists trained in the Modular Treatment condition also demonstrated significantly more positive attitudes toward EBTs following training than those trained in the standard EBT treatment condition (Borntrager, Chorpita, Higa-McMillan, & Weisz, 2009). Similarly, therapists who completed satisfaction measures at the completion of each case showed more significantly greater overall satisfaction with the protocol used in the modular treatment condition relative to the standard EBT treatment and UC conditions (Chorpita et al., 2015). These findings suggest that even if the clinical outcomes had been equivalent to those obtained using a standard EBT design, the modular design had important advantages for dissemination and sustainability, given that therapists showed greater inclination to continue using the approach with other cases beyond those involved in the study (Palinkas et al., 2013).

Given the advantages observed with this particular treatment design, affecting both clinical/functional outcomes and therapist implementation, and given the clear need to increase the efficiency and the reach of current effective interventions, we sought to test the performance of a modular protocol for youth in a more challenging context, serving a wider age range in an even more diverse participant sample and targeting an even wider range of clinical concerns. Thus, the current study used an updated version of the modular approach to therapy for children (MATCH) protocol, specifically adapted to include treatment procedures relevant to traumatic stress and child welfare involvement and whose materials were translated for Spanish speaking caregivers. The age

range was widened from 7 to 13 years to 5 to 15 years, and participants were selected from very low income urban settings in Los Angeles County and San Bernardino County. This current trial also sought to answer the question of how MATCH would perform relative to typical community-based implementation of EBTs, given that usual care services are increasingly characterized by training and implementation initiatives involving the use of EBTs.

Method

All study procedures were approved by the Institutional Review Board of the University of California, Los Angeles as well as by those institutional review boards of participating service agencies that requested independent reviews.

Measures

Brief problem checklist (BPC)—Child and caregiver versions. The BPC is a 12-item interview, with separate youth and caregiver forms, that was derived using item response theory and factor analysis with data obtained from the Child Behavior Checklist−Youth Self-Report (CBCL-YSR; Achenbach & Rescorla, 2001), using a sample of 2,332 youth (Chorpita et al., 2010). This brief measure was designed to be administered to assess change over time, using a frequent assessment schedule (e.g., weekly), with administration time averaging less than one minute. The BPC comprises three scales: Internalizing Problems, Externalizing Problems, and Total Problems. Chorpita and colleagues' (2010) study of 184 children found that the BPC yielded good internal consistency, test–retest reliability estimates, and large and significant correlations with paper-and-pencil administrations of the CBCL-YSR as well as with diagnoses obtained from a structured diagnostic interview. Longitudinal data gathered over 6 months of treatment demonstrated that the BPC significantly predicted change on related measures of child symptoms. Estimates obtained from random coefficient growth models have shown that the BPC yields higher slope reliabilities when given weekly relative to the CBCL-YSR given every 3 months. A Spanish translation of the BPC, derived using the translation procedures recommended by Brislin (1970), was administered to the predominantly Spanish-speaking caregivers. The BPC was administered at the baseline and posttreatment assessments as well as on a weekly basis during the course of treatment. Internal consistency at baseline administration for Spanish-speaking caregivers ($n = 84$) was comparable to that for English-speaking caregivers ($n = 112$): .83 versus .81, respectively (total score); .81 versus .84, respectively (Externalizing); and .82 versus .81, respectively (Internalizing).

Top problems assessment (TPA)—Child and caregiver versions. The TPA collects information about the problems of greatest clinical concern to youth and their caregivers. First, youth are asked to nominate problems that they are most concerned about. Then, the interviewer records youths' responses in their own words (e.g., "I'm scared that something bad will happen to my mom") and prompts youth to rate the severity of each problem on a 0 to 10 scale, with higher ratings indicating greater problem severity. Next, youth are given a list of the problems that they identified and are asked to select the one that is currently the

biggest problem, which is assigned Rank 1; the interviewer then asks youth to select the next biggest problem and assigns that Rank 2 and continues with this process until a ranked list of youth-nominated problems has been generated. The TPA is then completed with youths' caregivers to generate a list of caregiver-nominated problems. The TPA has evidenced acceptable test–retest reliability, convergent and discriminant validity, sensitivity to change, and slope reliability (Weisz et al., 2011). Weisz et al. (2011) noted that TPA-reported concerns often did not correspond to clinical symptom scales (between 38% and 80% of the time, depending on scales and informants), suggesting that the TPA is perhaps best characterized as a measure of functional impairment and consumer-relevant concerns in general. The TPA was administered at the baseline and posttreatment assessments; during the course of treatment, caregivers and youth were asked to provide severity ratings for their top three nominated problems on a weekly basis. For Spanish speaking caregivers ($n = 84$), the measure was administered in Spanish and demonstrated comparable internal consistency to that for English-speaking caregivers ($n = 112$; .76 versus .72, respectively).

Revised Child Anxiety and Depression Scales (RCADS)—Child and caregiver versions. The RCADS is a 47-item questionnaire, with separate forms for youth and caregivers, which assess for *DSM-IV*, American Psychiatric Association (2000) related anxiety and depression dimensions in children and adolescents. The measure prompts youth and their caregivers to rate how often the youth experience each item (e.g., "worries about looking foolish," "feels sad or empty") on a 4-point Likert scale ranging from 0 (*never*) to 3 (*always*). The RCADS produces six subscales: Generalized Anxiety Disorder, Separation Anxiety Disorder, Social Phobia, Panic Disorder, Obsessive−Compulsive Disorder, and Major Depressive Disorder. The measure also yields a Total Anxiety Scale (sum of the five anxiety subscales) and a Total Internalizing Scale (sum of all six subscales). Raw scores are calculated by summing the items that comprise each subscale, and *t* scores are calculated on the basis of youth's raw score, age, and gender (*t* scores of 65 or higher are considered to be in the borderline clinical range, and *t* scores of 70 or higher are considered to be in the clinical range). The RCADS has demonstrated acceptable internal consistency, convergent and divergent validity, and a factor structure that converges with the measure's six subscales in community (Chorpita et al., 2000; Ebesutani et al., 2010, 2011; Ebesutani, Tottenham, & Chorpita, 2015) and clinical samples (Chorpita et al., 2005) in the United States and Australia (de Ross et al., 2002). A Spanish translation of the RCADS has also evidenced strong psychometric properties (Park, Ebesutani, Bose, & Chorpita, 2015). The RCADS was administered at the baseline assessment.

University of California at Los Angeles Post-traumatic Stress Disorder Reaction Index (UCLA PTSD Index)—Child, adolescent, and caregiver versions. The UCLA PTSD Index (Steinberg, Brymer, Decker, & Pynoos, 2004), which has separate forms for children (ages 7 to 12), adolescents (ages 13 to 18), and their caregivers, screens for exposure to traumatic events and assesses posttraumatic stress disorder (PTSD) symptoms in children and adolescents. The measure is administered in four parts. Part 1 screens and categorizes any traumatic events that the youth has experienced and, if exposure to more than one event is endorsed, identifies the traumatic event that is currently the most

bothersome. Part 2 evaluates the presence or absence of feelings associated with exposure to a traumatic event. Part 3 assesses posttraumatic stress symptoms experienced during the past month, and Part 4 assesses how often the symptoms occurred over the past month. The UCLA PTSD Reaction Index uses a scoring algorithm to generate a total score. Total scores that fall at or above the cutoff of 38 have the greatest sensitivity and specificity for detecting PTSD. The UCLA PTSD Index is available in English and Spanish versions. The measure has evidenced good internal consistency as well as convergent and factorial validity (Steinberg et al., 2013), and internal consistency coefficients in this sample were .92 (English language) and .94 (Spanish language) at the baseline assessment.

Strength and Difficulties Questionnaire (SDQ)—Child and caregiver versions. The SDQ is a 25-item measure with separate forms for youth and their caregivers. Items assess youths' emotional and behavioral difficulties on a 3-point Likert scale, ranging from 0 (*not true*) to 2 (*very true*). The SDQ produces five subscales: Emotional Symptoms, Conduct Problems, Hyperactivity−Inattention, Peer Problems, and Prosocial. The SDQ also yields Internalizing Difficulties (sum of the Emotional Symptoms and Peer Problems subscales), Externalizing Difficulties (sum of the Conduct Problems and Hyperactivity−Inattention subscales), and Total Difficulties (sum of the Internalizing and Externalizing scales) scales. Raw scores on the caregiver version of the SDQ were considered to be in the high range at 17 or greater for Total Difficulties, at 5 or greater for Emotional Problems, at 4 or greater for Conduct Problems, at 8 or greater for Hyperactivity−Inattention, at 4 or greater for Peer Problems, and at 6 or less for Prosocial. Raw scores on the child version of the SDQ were considered to be in the high range at 15 or greater for Total Difficulties, at 5 or greater for Emotional Problems, at 4 or greater for Conduct Problems, at 6 or greater for Hyperactivity−Inattention, at 3 or greater for Peer Problems, and at 6 or less for Prosocial. The SDQ has been shown to have good factorial, internal consistency, test–retest reliability, and cross-informant reliability (Goodman, 2001). The Spanish version of the SDQ has similarly demonstrated good internal consistency and support for a five-factor structure (Rodríguez-Hernández et al., 2012). The SDQ was administered at the baseline and posttreatment assessments.

Services assessment for children and adolescents (SACA)− Parent version. The SACA-Parent version is a semistructured interview of lifetime and past-year use of 30 different service settings grouped into three categories: inpatient, outpatient, and school-based services. To examine youths' service use at more frequent intervals, this study used a simplified version of the SACA, focusing on utilization of only 23 different types of inpatient, outpatient, and school-based services and only over the last 3 months. The SACA has demonstrated favorable test–retest reliability (Horwitz et al., 2001), parent–child agreement (Stiffman et al., 2000), and concordance with service records (Hoagwood et al., 2000). The SACA was administered at the baseline and posttreatment assessments.

Services for children and adolescents−Parent interview (SCAPI). The SCAPI (Jensen et al., 2004) is a structured interview of service use that is grouped into 10 categories: child medication, monitoring of child medication, family therapy, counseling/therapy/in-home case management, overnight stays, special help in school, summer/weekend treatment program, parent

classes/groups, parent counseling for own difficulties, and parent medication. For the purposes of this study, the SCAPI was used to collect data that were distinct from that provided by the SACA and, accordingly, focused on factors related to youths' use of psychotropic medication—including name of medication, reason for prescription, dosage, and start and end dates. The SCAPI, particularly the child medication assessment, has demonstrated favorable psychometric properties (Hoagwood et al., 2000). The SCAPI was administered at the baseline and posttreatment assessments as well as during the course of therapy when caregivers endorsed a change in their youth's medication.

Client Satisfaction Questionnaire—Child and caregiver versions. The Client Satisfaction Questionnaire is an eight-item measure, with separate forms for youth and their caregivers, that assesses client satisfaction with the therapy services. The measure prompts youth and their caregivers to rate each item on a 4-point Likert scale ranging from 1 to 4, such that higher scores indicate greater satisfaction. The CSQ (in English and in Spanish) has demonstrated good internal consistency and moderate correlations with several service use variables (Attkisson & Zwick, 1982; Roberts, Atrkisson, & Mendias, 1984). The CSQ was administered at the posttreatment assessment.

Participants

Inclusion and exclusion criteria. Youth between the ages of 5 and 15 referred to their local public mental health agency were included if their primary clinical concerns involved anxiety, depression, disruptive behavior, or traumatic stress. Youths' primary clinical concerns were determined using the TPA procedure, which involved a consensus rating from youth and caregiver reports (Weisz et al., 2011) and scores on corresponding measures that demonstrated clinical elevations in any of the four areas (e.g., rating of "high" or above for the Conduct Problems Scale of the Strength and Difficulties Questionnaire for disruptive behavior concerns, a *t* score of 65 or higher on the relevant scale of the RCADS for anxiety or depression concerns, Chorpita et al., 2000, a score of 25 or higher on the UCLA PTSD Index for traumatic stress concerns). Youths were excluded if (a) they had evidence of intellectual disability, autism, or psychosis as identified in the initial assessment procedures; (b) they were 14 or older and had current juvenile justice system involvement; (c) they reported an attempted suicide within the past 3 months; or (d) they scored 38 or higher on the UCLA PTSD Index and had a primary clinical focus of traumatic stress (this criterion was intended to exclude youths for whom severe trauma was to be the primary focus of treatment, given the availability of evidence-based practices in the system for traumatic stress; however, youth with UCLA PTSD Index scores of 38 or higher were not excluded if the TPA procedure identified a problem area other than traumatic stress as the primary concern; only two youth were excluded for this reason). Figure 1 shows the flow of youth into the study according to CONSORT guidelines (Campbell, Elbourne, Altman, & CONSORT Group, 2004).

Sample demographics. Youths were ages 5 to 15 (N = 138; M = 9.30, SD = 2.82), slightly more than half (55.1%; n = 76) were boys, and the sample was ethnoracially diverse: 78.3% Latino/a, 10.1% African American, 8.0% Multiethnic, and 3.6% Caucasian. Most families reported living in poverty, with 91.3% of
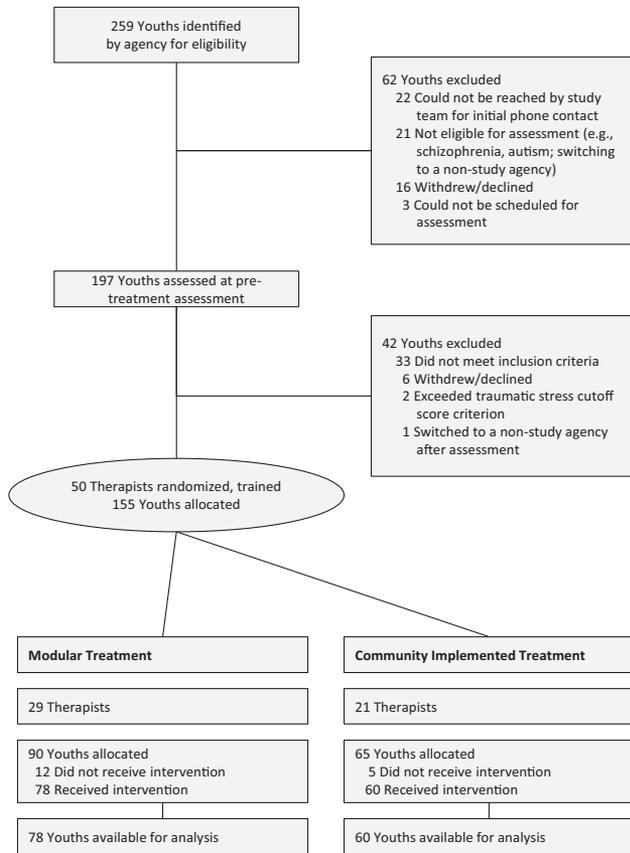
259 Youths identified by agency for eligibility

62 Youths excluded
22 Could not be reached by study team for initial phone contact
21 Not eligible for assessment (e.g., schizophrenia, autism; switching to a non-study agency)
16 Withdrew/declined
3 Could not be scheduled for assessment

197 Youths assessed at pre-treatment assessment

42 Youths excluded
33 Did not meet inclusion criteria
6 Withdrew/declined
2 Exceeded traumatic stress cutoff score criterion
1 Switched to a non-study agency after assessment

50 Therapists randomized, trained
155 Youths allocated

Modular Treatment

29 Therapists

90 Youths allocated
12 Did not receive intervention
78 Received intervention

78 Youths available for analysis

Community Implemented Treatment

21 Therapists

65 Youths allocated
5 Did not receive intervention
60 Received intervention

60 Youths available for analysis

Figure 1. CONSORT flow diagram of youths referred to project.

Table 1
Problems Identified at Baseline Using the Top Problems Assessment (N = 138)

Focus	Primary concern (%)	Any concern (%)
Conduct/disruptive behavior	60 (43.5)	110 (79.7)
Depression	39 (28.3)	101 (73.2)
Anxiety	38 (27.5)	118 (85.5)
Traumatic stress	1 (0.7)	24 (17.4)
Inattention/hyperactivity	0 (0)	46 (33.3)
Elimination disorder	0 (0)	6 (4.3)

2), on the UCLA PTSD Index, 79 caregivers (57.2%) reported that their youth experienced at least 1 traumatic event ($M = 1.31$; $SD = 1.62$; range $= 0-8$), and of the 105 youth who completed the youth version of the UCLA PTSD Index, 64 (61.0%) reported previously experiencing at least one traumatic event ($M = 1.31$; $SD = 1.56$; range $= 0-7$). In addition, a significant portion of the sample had child welfare involvement. Specifically, 10.0% of youth were registered with Los Angeles County Department of Child and Family Services (DCFS) for a case open during the study, with an additional 22.5% of the sample having a DCFS case that had closed prior to beginning participation in the study.

Table 2
Baseline Scores and Clinical Elevation Rates on Standardized Measures by Condition

Measure	MATCH			CIT		
	M	SD	Elevated	M	SD	Elevated
SDQ−Caregiver						
Total difficulties	18.99	6.33	82.1%	18.65	6.96	76.7%
Emotional problems	5.05	2.59	73.1%	4.80	2.77	65.0%
Conduct problems	4.26	1.93	79.5%	4.18	2.25	75.0%
Hyperactivity	6.44	2.65	66.7%	5.93	2.82	58.3%
Peer problems	3.24	2.02	60.3%	3.73	2.43	70.0%
Prosocial problems	6.41	2.05	82.1%	6.52	2.41	73.3%
SDQ−Youth						
Total difficulties	16.53	5.17	58.1%	18.10	6.40	68.3%
Emotional problems	3.87	2.65	25.8%	4.22	2.50	24.4%
Conduct problems	3.08	1.33	32.3%	3.61	2.34	41.5%
Hyperactivity	4.87	1.82	37.1%	5.10	1.96	39.0%
Peer problems	4.71	1.68	80.6%	5.17	1.92	87.8%
Prosocial problems	7.63	1.80	96.8%	7.73	1.72	97.6%
RCADS−Caregiver						
Total	60.39	16.40	51.2%	61.41	15.51	53.3%
Total anxiety	58.33	16.33	70.5%	59.10	15.30	43.3%
Depression	64.32	17.42	41.0%	66.02	16.40	78.3%
RCADS−Youth						
Total	49.38	14.41	22.2%	51.55	14.32	33.3%
Total anxiety	49.41	14.41	22.2%	50.88	13.98	23.8%
Depression	49.23	12.39	23.8%	53.28	14.28	35.7%
UCLA PTSD Index						
Caregiver	13.47	18.16	25.6%	13.88	17.28	23.3%
Youth	14.75	16.35	28.6%	7.86	13.42	9.5%

Note. Clinical elevations were calculated using published cutoff scores for each instrument. MATCH = modular approach to therapy for children; CIT = community implemented treatment; SDQ = Strengths and Difficulties Questionnaire; RCADS = Revised Child Anxiety and Depression Scale; PTSD = posttraumatic stress disorder; UCLA PTSD Index = University of California at Los Angeles Post-traumatic Stress Disorder Reaction Index.

the sample reporting household income of $39,000 or less and 71.0% reporting income of $19,000 or less, supporting an average number of 4.1 dependents per family ($SD = 1.5$). Nearly half of primary caregivers reported not having completed high school (47.8%). The majority of families (58.0%) were led by a single parent (24.6% never married, 19.6% separated, 11.6% divorced, and 2.2% widowed), and 40.6% had more than one adult caregiver (28.3% married and 12.3% unmarried living with partner). Two families (1.4%) did not report marital status. Nearly half of the participating families spoke Spanish as their primary language at home ($n = 64$; 46.38%), and the majority of families (68.84%) had at least one parent born outside the United States (with most of those from Mexico or El Salvador and the remainder from approximately 10 other Central and South American countries or the Middle East).

Table 1 shows the clinical focus for youth as determined by the TPA procedure. The majority of primary concerns nominated by families involved conduct problems, which were also the most common problem noted overall (even when not the primary concern). Depression and anxiety concerns represented somewhat less than half of the sample, and were present as a concern in the majority of the sample. Traumatic stress was nominated as a concern in less than one fifth of the sample; however, other data indicated that experience with trauma was more common than indicated by the TPA. For example, although youth traumatic stress was rarely the primary reason for seeking services (see Table

Table 2 shows baseline scores for measures used to determine eligibility for the study. Although significant elevations were required on only one corresponding scale for study eligibility, a high rate was observed for all of the scales used, suggesting that youth often had multiple problems or areas of impairment. This pattern was consistent with the TPA baseline assessments. No baseline measures differed significantly across study condition.

Setting

Therapists and service organizations. Fifty therapists in three different community agencies in the Southern California area delivered treatment. Therapists were 96% female, had an average of 3.30 years clinical experience post degree ($SD = 3.06$), and 40% described their professional specialty as social work, 44% marriage and family therapy, and 16% counseling or clinical psychology. All therapists had graduate training in their specialties, with 14% reporting a doctoral degree (i.e., PhD or PsyD) and 86% reporting a master's degree (e.g., MA, MS, or MSW) as their highest education obtained. Twenty-two percent held a state license. Therapists saw an average of 2.76 cases who participated in this study ($SD = 2.06$; range $= 1-9$). There were no significant differences across condition on any therapist demographic or professional experience characteristics, or on the number of study cases seen.

System context. This trial occurred primarily in the context of a large-scale implementation of EBTs in Los Angeles County (LAC). LAC is the largest county-operated mental health service system in the nation and supplies the county with almost 90% of its child and adolescent services (Southam-Gerow et al., 2014). In 2009, just prior to the outset of this trial, LAC Department of Mental Health began an initiative known as the Prevention and Early Intervention (PEI) transformation, which amended contracts with over 100 mental health service agencies to limit youth mental health service reimbursement to 52 county-approved EBTs (see Southam-Gerow et al., 2014, for a detailed description of the service context and PEI transformation). Thus, all participating agencies and their therapists were subject to considerable external pressure to implement EBTs with their entire caseloads. These requirements had two important implications for this trial. First, therapists who were in the community implemented treatment control condition of this trial were uniformly expected to be trained in and to provide one or more county-approved EBTs during the course of this investigation. Thus, the control condition did not represent the commonly encountered UC condition that is typical of many systems, allowing therapists relative autonomy to select what service models to use. At the same time, given that we did not assess the quality of implementation support, the control condition did not represent a formal EBT comparison condition, in which model-specific supports and fidelity assessments would be regularly provided by the research team. We thus conceptualized the control group as a community implementation of EBTs. The second implication of the PEI transformation was the potential impact on system and organizational climate (see Reding et al., 2016). For example, therapists repeatedly expressed concerns regarding whether the approved list of EBTs was fully relevant and appropriate, and whether training and support in their administration was feasible.

Experimental Design

This investigation used a cluster randomization design (Campbell et al., 2004; Donner & Klar, 2000; Fayers, Jordhøy, & Kaasa, 2002) with therapists assigned to condition using blocked randomization. The rationale of using a therapist as the cluster is that youths would be nested within therapist, and such a design would better protect against diffusion of the independent variable across cases. A computerized random number generator produced an unpredictable sequence of numbers representing condition, which were assigned to therapists. Block size was the entire cohort of participating therapists within each organization, and the allocation ratio for each block was 1:1. Allocation concealment was maintained through the use of therapist ID numbers. Youth who were referred to randomized therapists were routinely screened for eligibility and offered treatment if a study assessment confirmed eligibility as described above. Families were compensated for participating in telephone assessments with the university-based study team but were not compensated for participating in treatment services with their community therapists.

Treatment Conditions

Community implemented treatment (CIT) condition. Therapists randomized to CIT used treatment procedures as they normally would in the context of county-mandated use of EBTs, and therapy continued until normal client termination. County quality management procedures included widespread training in EBTs by their developers, ongoing audits, contract reviews, supervision, and reimbursement contingencies to support the use of these treatments. Youth allocated to the CIT condition began a new episode of care with CIT therapists. CIT therapists reported receiving workshop-based training and consultation in an average of 2.14 county-supported EBTs (range $= 0-6$), with 89.3% reporting having been fully trained in the use of at least one EBT. The proportion of therapists trained in each of the county's most common EBTs was as follows: trauma-focused cognitive-behavioral therapy (66.7%; e.g., Cohen & Mannarino, 1996), seeking safety (52.4%; e.g., Najavits, Gallop, & Weiss, 2006), Positive Parenting Program (28.6%; e.g., Sanders, Markie-Dadds, Turner, & Ralph, 2004), depression treatment quality improvement (19.0%; e.g., Asarnow et al., 2005), Incredible Years (19.0%; e.g., Webster-Stratton, 2005), parent–child interaction training (14.3%; e.g., Eyberg & Matarazzo, 1980), and Cognitive-Behavioral Intervention for Trauma in Schools (9.5%; e.g., Stein et al., 2003). Notably, there were no EBTs in the county-supported array specifically targeting anxiety; whereas three of the most widely trained treatments targeted traumatic stress, three targeted disruptive behavior, and one targeted depression. County-sponsored training in these EBTs began approximately 4 years prior to the current study, with trainings and boosters continually available throughout the study period. The majority of therapist trainings, including boosters, occurred prior to the study onset (64%), with 57% occurring within a year of the study onset for these therapists.

MATCH treatment condition. Therapists in this condition used the Modular Approach to Therapy for Children (MATCH; Chorpita & Weisz, 2005; Chorpita & Weisz, 2009), a collection of 33 modules that correspond to cognitive–behavioral and parent-management training procedures common among many EBTs targeting youth anxiety, depression, conduct problems, and trau-

matic stress (e.g., exposure, problem solving, rewards, time out, cognitive restructuring, relaxation). The protocol uses guiding algorithms that default to a sequence that is similar to many EBTs targeting these four areas, but that also allow for real-time adaptation to address any interference if outcomes measures gathered each week demonstrate poor treatment response (see Chorpita, Bernstein, Daleiden, & Research Network on Youth Mental Health, 2008). Specific strategies regarding how to address impediments to treatment implementation are collaboratively determined by the treatment team (including the research team members), such that the course of action arises from a collaborative review of the strategies outlined in the manual, recommendations by the MATCH consultant, therapist knowledge about the case, and feedback on outcomes and practices covered so far. Thus, for example, MATCH therapists could use procedures earlier than indicated by the manual, by jumping ahead in the protocol, could omit procedures that did not seem well-suited, or could use procedures for multiple problem areas concurrently within a single treatment episode. This collaborative adaption of the protocol in the context of client information and consultant input was also used to guide the age-specific adaptations needed for such a wide age range in this study. For example, worksheets and exercises could be adapted, omitted, or replaced with functionally similar ones that best fit the clients' developmental level, based on a structured, data-driven collaboration in the consultation process (Chorpita & Daleiden, 2014).

Observational coding of 62 digital recordings randomly selected from 2,245 sessions showed that 78% of sessions from the MATCH condition used content consistent with MATCH procedures as compared with 7% of sessions from the CIT condition, $\chi^2(1) = 32.16$, $p < .001$. Treatment sessions were coded using the same coding procedures as those described in Weisz et al. (2012). Coders consisted of two doctoral level coders and one doctoral psychology graduate student—all of who demonstrated acceptable intercoder reliability (intraclass correlation coefficient [ICC] ≥ .70) prior to coding study sessions. TPA scores (Weisz et al., 2011) were used to determine the initial focus of the intervention for MATCH.

The MATCH condition also involved the use of an ongoing measurement system, the MATCH dashboards, which provide a plot of outcome measures (e.g., internalizing symptoms) as well as a temporally aligned "event plot" that represents MATCH practices delivered over time (see Chorpita et al., 2008). Dashboards are populated with data gathered by the research team from the youth, family, and other sources, and are reviewed in weekly consultation meetings with therapists to help them make decisions about treatment planning, practice delivery, and adaptation at future treatment sessions based on the treatment team's integrative interpretation of practices delivered and progress observed.

MATCH Consultants

This study included five postdoctoral consultants who received training in the MATCH protocol and participated in weekly clinical meetings with the treatment developer. All consultants were female and were doctorate-level psychologists. They had an average age of 31.60 years (SD = 2.07), and their race/ethnicity was reported as 40% Caucasian, 40% Asian, and 20% Latina. Consultants had a baseline average of 7.00 years of postundergraduate

professional and clinical experience (SD = 1.22) and 0.80 years of experience providing clinical supervision (SD = 0.84). Prior to this study, consultants had experience working on an average of 7.20 research projects (SD = 3.35). Eighty percent reported their primary theoretical orientation as cognitive-behavioral and 20% as eclectic. Prior to the study, consultants reported having received training in three to seven EBTs (M = 4.80, SD = 1.48).

Therapist Training and Consultation

Therapists in the MATCH condition were trained together for 5 days by the treatment developer and the MATCH consultants, followed by individual weekly consultation from the consultants during the course of the trial. Consultants in turn participated in weekly discussions with the treatment developer that incorporated review of measurement feedback on client progress and practice history (Chorpita et al., 2008). Aside from their training in county-supported EBTs noted above, CIT therapists received only the relevant and routine supervision procedures in their local settings, with no intervention from project personnel, other than to retrieve audiotapes of CIT treatment sessions and session dates on a periodic basis.

Data Analysis

As in the initial evaluation with MATCH, all youth outcomes were obtained by assessors blind to treatment condition, and outcomes were evaluated longitudinally, given that treatment did not have a fixed length, thus rendering prepost analyses difficult to interpret. We estimated mixed effects regression models (Bryk & Raudenbush, 1992) with the following predictors: intercept, treatment condition, time (days since intake), and treatment condition by time. Unlike in the initial Child STEPs multisite trial (Weisz et al., 2012), a log-linear model did not fit the data well, with change functions often varying widely and showing no uniform pattern across participants (cf. Tang & DeRubeis, 1999). We thus estimated overall linear rate of change for all models, to determine which youths simply showed a faster overall rate of improvement.

We considered modeling up to five levels of nesting: repeated measures within informant within case within therapist within service organization. When we examined organization and therapist, however, boundary conditions were frequent (i.e., the highest level variance was small and poorly estimated). For the nonboundary three-level models with repeated measurements nested within individuals nested within therapists, therapist accounted for 1% to 4% of the variance in outcomes. For the three nonboundary three-level models with repeated measurements nested within individuals nested within organizations, organization accounted for 1% of the variance in outcomes. We then examined several four-level models with measures nested within informant within case within therapist. All of these resulted in boundary conditions with therapist variance estimated at zero; thus, for the final study analyses, we evaluated only three-level models, in which intercept and time were modeled as random effects at the informant level; the intercept only was modeled as random at the case level.

Given that youth aged 5 and 6 were unable to provide self-report data for the study, our preference was to estimate models using all available data for each participant when possible. Thus we explicitly tested whether youth and parent information was empirically

distinguishable (Kenny, Kashy, & Cook, 2006) in terms of fixed effects and variance components at the informant level. Specifically, for fixed effects, we examined interactions between a binary informant predictor and treatment condition, time, and the treatment condition by time interaction, retaining any statistically significant interactions and the necessary lower order effects to support them. Once we determined which fixed effects could be combined across informants, we conducted likelihood ratio tests to compare nested models that estimated either a single set of variance components at the informant level or separate sets of variance components for parents and youth, retaining the latter only if it resulted in a significant improvement in model fit. Thus, our preference was to use all data in a single model with informant as a predictor when this approach fit the data best.

Results

Symptom Outcomes

Externalizing. We first examined BPC Externalizing scores reported by youths and caregivers in a single three-level mixed effects regression analysis. The model for this particular outcome included a marginally significant rater by time interaction ($b = -0.0022$, $SE = 0.0011$, $t = -1.94$, $p = .055$), such that the rate of decrease in Externalizing scores was greater for youth than for caregivers; this effect did not differ by treatment condition. This model also estimated intercept and time trend variances separately for youths and caregivers, as preliminary analyses revealed marginally greater variance, $\chi^2(3) = 6.4$, $p = .09$, among caregivers. Table 3 shows the condition by time interaction, such that the rate of decrease in Externalizing scores was significantly greater for MATCH compared with CIT. Table 3 also shows the effect size for linear rates of change, calculated as the estimated difference in rate of change (i.e., the condition by time estimate) divided by the square root of the estimated time trend variance (Feingold, 2009), analogous to a Cohen's d effect (e.g., .20 = small, .50 = medium). Because of the difference in slope variability, effect sizes were computed separately by rater: .38 for caregivers and .50 for youths.

Internalizing. We then examined BPC Internalizing scores reported by youths and caregivers in a single three-level mixed

effects regression analysis. No significant rater interactions emerged for this particular outcome variable, nor were there any significant differences in variance for youths and caregivers. The effect of primary interest was the comparison of MATCH with the CIT control condition in terms of the linear decrease in scores over time. Table 3 shows the estimated slope for the condition by time interaction; the negative sign of this effect indicates that the rate of decrease in Internalizing scores was significantly greater for MATCH compared with CIT. This particular effect was associated with an effect size of .46.

Total symptoms. BPC Total Scores reported by youth and caregivers were similarly examined in a single three-level mixed effects regression analysis. No rater interactions were necessary to adequately model this particular outcome, nor were there any significant differences in variance for youth and caregivers. Table 3 shows the condition by time interaction, such that the rate of decrease in Total Scores was significantly greater for MATCH compared with CIT, with an effect size of .51.

To aid in overall interpretation of clinical improvement rates, we also calculated the percentage of youth who scored below the clinical threshold for caregiver-reported SDQ Total Difficulties Scale at posttreatment (using an intent to treat model whereby clients who did not provide data at posttreatment were considered not improved). This was a conservative comparative estimate of outcomes that ignored rate of change, given that treatment length was variable and could theoretically continue until outcomes were achieved in either condition. Nevertheless, at posttreatment, 46 of 78 youth (60%) were clinically improved in the MATCH condition, compared with 22 of 60 (36.7%) in the CIT condition. These differences in percentage improved were significantly different, $\chi^2(1) = 6.75$, $p < .01$.

Functional Outcomes

To investigate functional outcomes, we examined data from the TPA gathered weekly during the active phase of treatment, using a single three-level mixed effects regression analysis. The model for this particular outcome included a significant condition by time interaction ($b = -1.14$, $SE = 0.53$, $t = -2.16$, $p < .05$), such that MATCH youth reported higher Top Problems scores than CIT youth at study entry, while no such difference emerged for caregiver reports (note: these differences were noted for fitted model intercepts, but were not significant for baseline observed scores on the youth reported TPA: $M_{MATCH} = 4.08$, $SD_{MATCH} = 2.99$; $M_{CIT} = 4.51$, $SD_{CIT} = 2.95$; $t_{(98)} = 0.71$, $p = .48$). This model also estimated intercept and time trend variability separately for youth and caregivers, as preliminary analyses indicated significantly greater variance, $\chi^2(3) = 12.5$, $p < .01$, among youth. Table 3 shows the condition by time interaction, such that the rate of decrease in Top Problems scores was significantly greater for youth in the MATCH condition compared with CIT condition. Effect sizes were .56 for caregivers and .38 for youths. Although it is possible that the difference in rate of decrease among youths might be attributable in some part to differences between MATCH and CIT participants at study entry, this was not the case for caregivers, and we note that the condition by time interaction remained significant when analyses were conducted separately by reporter ($b = -0.0039$, $SE = 0.0018$, $t = -2.18$, $p < .05$, effect

Table 3

Coefficient Estimates for Condition by Time on Clinical and Functional Outcomes (N = 138 for Each Analysis)

Measure	b	SE	t	p	ES
Externalizing	−.00281	.00144	−2.46	.0155	.38/.50
Internalizing	−.00424	.00149	−2.85	.0052	.46
Total	−.00765	.00244	−3.14	.0022	.51
Top problems	−.00457	.0015	−3.05	.0028	.56/.38

Note. A negative estimate indicates that the modular approach to therapy for children (i.e., MATCH) treatment condition showed faster reduction in problem severity over time than did the control condition. When two effect sizes (ES) are reported, the first is for caregiver report and the second is for youth report, calculated separately because of significant Condition × Time × Informant interactions. All tests are statistically significant at $p <$ .05. $b =$ Condition × Time (day) interaction parameter, adjusted for all other effects in the model.

size = .47 for caregivers; $b = -0.0063$, $SE = 0.0028$, $t = -2.26$, $p < .05$, effect size = .54 for youths).

Engagement Outcomes

Session consistency. To estimate the regularity of attendance by condition, we examined the total number of documented treatment sessions within each condition, and divided the total length of service by the number of session. This yields an average number of days between sessions. Results showed that contact was slightly more frequent in the MATCH condition (average days between sessions = 8.51; $SD = 3.23$) than in the CIT condition (average days between sessions = 9.79; $SD = 3.99$). This difference was statistically significant, $t(136) = 2.08$, $p < .05$; Cohen's $d = 0.36$.

Termination status. We then tested for differences in termination status as a measure of treatment engagement. At final session, cases were coded as to whether they involved a *routine termination*, defined as the family, study treatment team, and consultants agreeing that stopping treatment was appropriate because goals were met or that no further gains would be expected from the current treatment. Reasons for not being classified as routine termination included families moving away, actively quitting or withdrawing from service, or becoming lost to the study treatment team and agency. Nine cases were excluded from this analysis, because their services had ended as a result of a therapist-triggered event (e.g., therapist leaving job, taking maternity leave). For the remaining 129 cases, 34 (48%) were classified as having a routine termination in the MATCH condition, compared with 25 (43%) in the CIT condition. These rates of routine termination status were not significantly different across condition ($\chi^2(1) = 0.29$; $p = .59$).

Utilization Outcomes

We first examined whether there were significant differences in the total amount of study treatment sessions provided to the youth participants. Youth in the MATCH condition received an average of 21.65 treatment sessions ($SD = 15.60$), whereas youth in the CIT condition received an average of 30.22 treatment sessions ($SD = 21.61$). These differences were statistically significant, $t(136) = 2.68$, $p < .01$. The average length of treatment was 191.78 days ($SD = 137.97$) for the MATCH condition and 269.98 days ($SD = 181.52$) for the CIT Condition. These differences were also significantly different, $t(136) = 2.87$, $p < .01$.

We then investigated the amount of additional, nonstudy services that youth concurrently used during treatment. Table 4 summarizes the rates at which caregivers reported accessing other mental health services for their youth. Results showed that rates of service utilization were not significantly different across condition at the baseline assessment, $\chi^2(1) = .14$, $p = .71$. During the treatment phase of the trial, however, there were significantly fewer youth who supplemented their study treatment with another mental health service in the MATCH condition compared with the CIT condition, $\chi^2(1) = 6.96$, $p < .01$.

Relatedly, rates of medication use were not significantly different across condition at the baseline assessment ($p = .62$)—with three youth from the MATCH condition and two youth from the CIT condition using psychiatric medications. During the course of treatment, 7 (9%) youth in the MATCH condition and 12 (20%) youth in the CIT condition initiated or increased their medication usage. These rates of initiating or increasing medication use did not differ significantly across conditions, $\chi^2(1) = 3.47$; $p = .06$. However, youth in the MATCH condition used significantly fewer types of medications (e.g., antidepressants, stimulants; $M = .13$, $SD = .37$) relative to those in the CIT condition ($M = .30$, $SD = .67$), controlling for baseline medication use, $F(1, 135) = 4.72$, $p < .05$.

Consumer Outcomes

To assess consumer satisfaction, we examined scores on the Client Satisfaction Questionnaire, Youth, and Caregiver Reports. Scores in both conditions were almost uniformly high. Caregiver average scores were 28.9 (out of a maximum 32) for MATCH ($SD = 3.70$) and 28.21 for CIT ($SD = 4.24$). These scores did not differ significantly, $t(136) = 1.03$, $p = .30$. Scores on the youth report were similarly high for both conditions (MATCH $M = 27.6$, $SD = 3.41$; CIT $M = 28.6$, $SD = 4.15$). These were also not significantly different, $t_{(136)} = 1.55$, $p = .12$.

Discussion

These findings add to the support for MATCH as an efficient model for use in community mental health contexts. MATCH showed superior rates of improvement over time on both clinical and functional outcomes, compared with routine community implementation of multiple evidence-based practices. These findings replicate those from the initial Child STEPs multisite trial, which found significantly faster improvement for youth receiving MATCH as compared with three standard EBTs and a usual care control group (Weisz et al., 2012). MATCH in the current trial was also more efficient than the alternative, demonstrating a significantly shorter treatment duration than the CIT control group, a finding that was consistent with those reported by Weisz et al. (2012). In the Wesiz et al. study, length of treatment for MATCH (210.15 days) and UC (275.49 days) was also similar to that observed the current study (191.78 vs. 269.98 days respectively), which in turn was consistent with Los Angeles County averages as

Table 4

Percentage of Youth Receiving Additional Services by Condition at Baseline and During Treatment

Service setting	MATCH ($n = 78$)		CIT ($n = 60$)		χ^2	p
	n	%	n	%		
Baseline						
Inpatient	3	4	1	2	.572	.449
Outpatient	4	5	6	10	1.198	.274
School	4	5	3	5	.001	.973
Any of the above	10	13	9	15	.136	.713
During treatment						
Inpatient	0	0	1	2	1.309	.252
Outpatient	2	3	8	13	5.852	.016*
School	3	4	5	8	1.250	.263
Any of the above	5	6	13	22	6.959	.008*

Note. MATCH = modular approach to therapy for children; CIT = community implemented treatment.
* $p < .05$.

well as other research on community mental health for children (e.g., Garland et al., 2010). Moreover, although the initial Child STEPs multisite trial did not report significantly fewer treatment sessions for MATCH, the current trial did find such differences as compared with the CIT control. Finally, consistency of sessions in the current trial (average of 8.51 days between sessions) was greater than the average consistency across the EBT conditions reported in the multisite trial (11.96 days).

The current findings are consistent with those reported by Park et al. (2016), regarding observed advantages for MATCH in reducing use of other mental health services. New findings in the current trial also showed that MATCH was associated with lower rates of initiating or increasing psychotropic medication (although these were not significant—because of the low number of youth in this category—the effect was large), as well as significantly lower rates of the use of different types of medications (although a limitation in this study is that the psychometric properties of both the medication and service use measures [i.e., the SCAPI and SACA] have not been established for their Spanish-language versions). Given the involvement of many study participants in the child welfare system, these findings are particularly relevant to the concerns recently raised by the Administration for Children and Families at the national level about overuse of psychotropic medication and the need for strategies to reduce medication use in child welfare systems (e.g., Raghavan et al., 2005; U.S. Department of Health and Human Services, Administration on Children, Youth & Families, 2012).

The current trial also extends the scope of use of MATCH relative to the initial multisite trial, given that youth in the current study experienced greater levels of poverty, lower rates of speaking English in the home, a wider age range (5 to 15 years compared with 7 to 13 years), and the ability to target traumatic stress through new protocol content. Along those lines, the ability to extend the use of MATCH to focus on traumatic stress required minimal revision to the original protocol, other than to add two new treatment modules and a coordinating flowchart specific to trauma. Thus, the ability to expand MATCH content with relative ease highlights the efficiency of updating modular treatment in design time (e.g., between trials; Chorpita & Daleiden, 2014), as proposed by Chorpita, Daleiden, and Weisz, (2005).

A notable limitation of these findings involves the lack of study cases for which trauma was selected as the primary treatment focus. That is, although traumatic stress was part of the clinical profile of many cases, this trial did not allow for a test of whether MATCH was effective for cases having traumatic stress as the primary focus. Further research is needed to determine whether MATCH is in fact effective for cases with a primary focus of trauma. Nevertheless, the current findings suggest two important points: (a) In a large community sample with high rates of traumatic stress, families almost uniformly prioritized a treatment focus on other mental health concerns, and (b) MATCH appeared to perform well in a context in which traumatic stress was highly prevalent and thus a trauma-informed treatment could be warranted.

Another important point concerns the interpretation of the primary findings. Unlike the Child STEPs multisite trial, the current study was not a controlled test of alternative designs for EBT, in which use of regular consultation, measurement feedback, and protocol content were all held equivalent and the two EBT condi-

tions differed primarily in terms of arrangement of practice content and decision-making processes. In the current trial, there was regular consultation and use of measurement feedback in the MATCH condition, as such features are inherently part of the MATCH protocol itself. However, there was no required use of measurement feedback in the CIT condition, and there is reason to believe that programmatic consultation was less consistent in the CIT condition. Thus, in this study, it is possible that differences observed were due to differences in the amount of model-specific supervision and outcome feedback provided. Nevertheless, previous research with MATCH that controlled for a higher dose of supervision and feedback did not find any effects for those features when used with three evidence-based treatments other than MATCH (Chorpita et al., 2013; Weisz et al., 2012). Although it may be of concern as to whether such use of feedback and supervision would persist without developer involvement, there is evidence that related models using highly similar procedures, such as Managing and Adapting Practice (Chorpita & Daleiden, 2014), have had enduring and widespread success with implementation outside the context of research trials (e.g., Southam-Gerow et al., 2014).

Another possible explanation for the current findings could involve the lack of a specific treatment in the CIT condition targeting anxiety, which represented the primary concern of over a quarter of the study participants; however, this lack of a target-specific protocol within CIT does not explain the differences observed on measures of disruptive behavior, for instance, so could not fully account for the observed differences. Moreover, the CIT array was determined by county leadership prior to the study and thus represents a naturalistic implementation of evidence based practices without the input of a specific research team, which can be subject to both gaps and inefficiencies (see Chorpita, Bernstein, & Daleiden, 2011).

A related issue involves our finding that only 7% of the CIT content was consistent with MATCH content. At first glance, this might seem to indicate that the CIT condition had very little EBT content, which could be another reason for the differences observed. However, our coding of digital recordings was to determine specification of the independent variables, and was not a measure of EBT content, per se. In that light, coders noted that procedures from EBTs were commonly observed among all recordings (coders were blind to condition), but many such procedures were not specific enough to the MATCH codes to be endorsed, which suggests that although there was little diffusion of the independent variable (i.e., MATCH was not picked up spontaneously within organizations by control therapists), 7% is likely an underestimate of the amount of EBT practices that were being delivered in the CIT condition. Future research using integrity instruments for 6 to 10 of the most common CIT treatments might yield a better estimate of EBT implementation in that condition.

Finally, given prior observations by Palinkas et al. (2013) about the tendency for therapists trained in EBTs to subsequently implement them in an individualized fashion, it is also quite likely that the CIT condition had levels of flexibility comparable with or higher than those in the MATCH condition. Thus, the most reasonable interpretation of the current findings is that in a community mental health context, the use of a single structured, modular, transdiagnostic treatment approach that includes regular measurement feedback in consultation may have advantages over training

therapists in multiple EBTs and then relying on traditional supervision models that do not incorporate ongoing progress and practice monitoring.

Considering the collective findings regarding MATCH, several questions remain for future research. For example, MATCH is simply one working prototype of a modular architecture, constrained in its original design to include only those practices corresponding to three EBTs chosen for a control condition of its initial randomized controlled trial. Were MATCH to be built again "from the ground up," the selection of practice content could involve a more comprehensive review of all evidence-based practices and their practice elements (e.g., Bernstein, Chorpita, Daleiden, Ebesutani, & Rosenblatt, 2015), perhaps leading to superior or more efficient content. Future research should therefore continue to investigate alternative content in modular treatment systems to explore the advantages of specific content (e.g., are some procedures essential to obtain positive outcomes?), whether such content could be developed and contributed independently as part of a multideveloper network (e.g., can laboratories with independent expertise contribute specific modules into a single coordinated system?), and questions regarding overall protocol organization (e.g., is it better to have more or fewer modules in the system?).

One final point concerns the repeated observation that despite superior outcomes in this context, MATCH continues to show less than optimal levels of youth and family engagement. Thus, it may be important to investigate whether use and outcomes could be improved further by incorporating modules with a specific focus on treatment engagement (e.g., Becker et al., 2015) in addition to the engagement practices currently incorporated in MATCH. Until then, however, the primary clinical implication is that implementation of MATCH in similar systems seems like a productive tactic in the service of youth mental health system improvement efforts, given that it been shown to be effective and efficient with multiple treatment foci across diverse populations of youth and families and that it is well-liked by the therapists who use it.

References

Aarons, G. A., & Sawitzky, A. C. (2006). Organizational culture and climate and mental health provider attitudes toward evidence-based practice. *Psychological Services, 3,* 61–72.

Achenbach, T. M., & Rescorla, L. (2001). *Manual for the ASEBA School-Age Forms & Profiles*. Burlington, VT: University of Vermont, Department of Psychiatry.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Asarnow, J. R., Jaycox, L. H., Duan, N., LaBorde, A. P., Rea, M. M., Murray, P., . . . Wells, K. B. (2005). Effectiveness of a quality improvement intervention for adolescent depression in primary care clinics: A randomized controlled trial. *Journal of the American Medical Association, 293,* 311–319. http://dx.doi.org/10.1001/jama.293.3.311

Attkisson, C. C., & Zwick, R. (1982). The Client Satisfaction Questionnaire. Psychometric properties and correlations with service utilization and psychotherapy outcome. *Evaluation and Program Planning, 5,* 233–237. http://dx.doi.org/10.1016/0149-7189(82)90074-X

Becker, K. D., Lee, B. R., Daleiden, E. L., Lindsey, M., Brandt, N. E., & Chorpita, B. F. (2015). The common elements of engagement in children's mental health services: Which elements for which outcomes? *Journal of Clinical Child and Adolescent Psychology, 44,* 30–43. http://dx.doi.org/10.1080/15374416.2013.814543

Bernstein, A., Chorpita, B. F., Daleiden, E. L., Ebesutani, C. K., & Rosenblatt, A. (2015). Building an evidence-informed service array: Considering evidence-based programs as well as their practice elements. *Journal of Consulting and Clinical Psychology, 83,* 1085–1096. http://dx.doi.org/10.1037/ccp0000029

Bickman, L. (2008). A measurement feedback system (MFS) is necessary to improve mental health outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry, 47,* 1114–1119. http://dx.doi.org/10.1097/CHI.0b013e3181825af8

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,* 677–681. http://dx.doi.org/10.1176/ps.2009.60.5.677

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology, 1,* 185–216. http://dx.doi.org/10.1177/135910457000100301

Bryk, A. S., & Raudenbush, S. K. (1992). *Hierarchical linear models*. Newbury Park, CA: Sage.

Campbell, M. K., Elbourne, D. R., & Altman, D. G., & the CONSORT group. (2004). CONSORT statement: Extension to cluster randomised trials. *British Medical Journal (Clinical Research Ed.), 328,* 702–708. http://dx.doi.org/10.1136/bmj.328.7441.702

Chorpita, B. F., Bernstein, A., & Daleiden, E. L. (2011). Empirically guided coordination of multiple evidence-based treatments: An illustration of relevance mapping in children's mental health services. *Journal of Consulting and Clinical Psychology, 79,* 470–480. http://dx.doi.org/10.1037/a0023982

Chorpita, B. F., Bernstein, A., & Daleiden, E. L., & the Research Network on Youth Mental Health. (2008). Driving with roadmaps and dashboards: Using information resources to structure the decision models in service organizations. *Administration and Policy in Mental Health and Mental Health Services Research, 35*(1–2), 114–123. http://dx.doi.org/10.1007/s10488-007-0151-x

Chorpita, B. F., & Daleiden, E. L. (2014). Structuring the collaboration of science and service in pursuit of a shared vision. *Journal of Clinical Child and Adolescent Psychology, 43,* 323–338. http://dx.doi.org/10.1080/15374416.2013.828297

Chorpita, B. F., Daleiden, E. L., Ebesutani, C., Young, J., Becker, K. D., Nakamura, B. J., . . . Starace, N. (2011). Evidence-based treatments for children and adolescents: An updated review of indicators of efficacy and effectiveness. *Clinical Psychology: Science and Practice, 18,* 154–172.

Chorpita, B. F., Daleiden, E. L., & Weisz, J. R. (2005). Modularity in the design and application of therapeutic interventions. *Applied & Preventive Psychology, 11,* 141–156. http://dx.doi.org/10.1016/j.appsy.2005.05.002

Chorpita, B. F., Moffitt, C. E., & Gray, J. (2005). Psychometric properties of the Revised Child Anxiety and Depression Scale in a clinical sample. *Behaviour Research and Therapy, 43,* 309–322. http://dx.doi.org/10.1016/j.brat.2004.02.004

Chorpita, B. F., Park, A., Tsai, K., Korathu-Larson, P., Higa-McMillan, C. K., Nakamura, B. J., . . . the Research Network on Youth Mental Health. (2015). Balancing effectiveness with responsiveness: Therapist satisfaction across different treatment designs in the Child STEPs randomized effectiveness trial. *Journal of Consulting and Clinical Psychology, 83,* 709–718. http://dx.doi.org/10.1037/a0039301

Chorpita, B. F., Reise, S., Weisz, J. R., Grubbs, K., Becker, K. D., Krull, J. L., & the Research Network on Youth Mental Health. (2010). Evaluation of the Brief Problem Checklist: Child and caregiver interviews to measure clinical progress. *Journal of Consulting and Clinical Psychology, 78,* 526–536. http://dx.doi.org/10.1037/a0019602

Chorpita, B. F., & Weisz, J. R. (2005). *Modular approach to therapy for children with anxiety, depression, or conduct prblems*. Boston, MA: Harvard Medical School.

Chorpita, B. F., & Weisz, J. R. (2009). *MATCH-ADTC: Modular Approach to Therapy for Children with Anxiety, Depression, Trauma, or Conduct Problems*. Satellite Beach, FL: PracticeWise.

Chorpita, B. F., Weisz, J. R., Daleiden, E. L., Schoenwald, S. K., Palinkas, L. A., Miranda, J., . . . Research Network on Youth Mental Health. (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,* 999–1009.

Chorpita, B. F., Yim, L., Moffitt, C., Umemoto, L. A., & Francis, S. E. (2000). Assessment of symptoms of *DSM–IV* anxiety and depression in children: A Revised Child Anxiety and Depression Scale. *Behaviour Research and Therapy, 38,* 835–855. http://dx.doi.org/10.1016/S0005-7967(99)00130-8

Cohen, J. A., & Mannarino, A. P. (1996). A treatment outcome study for sexually abused preschool children: Initial findings. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 42–50. http://dx.doi.org/10.1097/00004583-199601000-00011

de Ross, R., Gullone, E., & Chorpita, B. F. (2002). The Revised Child Anxiety and Depression Scale: A psychometric investigation with Australian youth. *Behaviour Change, 19,* 90–101. http://dx.doi.org/10.1375/bech.19.2.90

Donner, A., & Klar, N. (2000). *Design and analysis of cluster randomization trials in health research*. London, England: Arnold.

Eaton Hoagwood, K., Jensen, P. S., Arnold, L. E., Roper, M., Severe, J., Odbert, C., . . . the MTA Cooperative Group. (2004). Reliability of the services for children and adolescents-parent interview. *Journal of the American Academy of Child and Adolescent Psychiatry, 43,* 1345–1354. http://dx.doi.org/10.1097/01.chi.0000139558.54948.1f

Ebesutani, C., Bernstein, A., Nakamura, B. J., Chorpita, B. F., & Weisz, J. R., & the Research Network on Youth Mental Health. (2010). A psychometric analysis of the Revised Child Anxiety and Depression Scale—Parent version in a clinical sample. *Journal of Abnormal Child Psychology, 38,* 249–260. http://dx.doi.org/10.1007/s10802-009-9363-8

Ebesutani, C., Chorpita, B. F., Higa-McMillan, C. K., Nakamura, B. J., Regan, J., & Lynch, R. E. (2011). A psychometric analysis of the Revised Child Anxiety and Depression Scales—Parent version in a school sample. *Journal of Abnormal Child Psychology, 39,* 173–185. http://dx.doi.org/10.1007/s10802-010-9460-8

Ebesutani, C., Tottenham, N., & Chorpita, B. (2015). The Revised Child Anxiety and Depression Scale—Parent version: Extended applicability and validity for use with younger youth and children with histories of early-life caregiver neglect. *Journal of Psychopathology and Behavioral Assessment, 37,* 705–718. http://search.proquest.com/docview/1756078864?accountid=14512. http://dx.doi.org/10.1007/s10862-015-9494-x

Eyberg, S. M., & Matarazzo, R. G. (1980). Training parents as therapists: A comparison between individual parent−child interaction training and parent group didactic training. *Journal of Clinical Psychology, 36,* 492–499. http://dx.doi.org/10.1002/jclp.6120360218

Fayers, P. M., Jordhøy, M. S., & Kaasa, S. (2002). Cluster-randomized trials. *Palliative Medicine, 16,* 69–70. http://dx.doi.org/10.1191/0269216302pm503xx

Feingold, A. (2009). Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychological Methods, 14,* 43–53. http://dx.doi.org/10.1037/a0014699

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,* 788–795. http://dx.doi.org/10.1176/ps.2010.61.8.788

Glisson, C., Landsverk, J., Schoenwald, S., Kelleher, K., Hoagwood, K. E., Mayberg, S., . . . the Research Network on Youth Mental Health. (2008). Assessing the organizational social context (OSC) of mental health services: Implications for research and practice. *Administration and Policy in Mental Health and Mental Health Services Research, 35*(1–2), 98–113. http://dx.doi.org/10.1007/s10488-007-0148-5

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry, 40,* 1337–1345. http://dx.doi.org/10.1097/00004583-200111000-00015

Henggeler, S. W., Schoenwald, S. K., & Pickrel, S. G. (1995). Multisystemic therapy: Bridging the gap between university- and community-based treatment. *Journal of Consulting and Clinical Psychology, 63,* 709–717. http://dx.doi.org/10.1037/0022-006X.63.5.709

Hoagwood, K., Horwitz, S., Stiffman, A., Weisz, J., Bean, D., Rae, D., . . . Leaf, P. (2000). Concordance between parent reports of children's mental health services and service records: The Services Assessment for Children and Adolescents (SACA). *Journal of Child and Family Studies, 9,* 315–331. http://dx.doi.org/10.1023/A:1026492423273

Horwitz, S. M., Hoagwood, K., Stiffman, A. R., Summerfeld, T., Weisz, J. R., Costello, E. J., . . . Norquist, G. (2001). Reliability of the services assessment for children and adolescents. *Psychiatric Services, 52,* 1088–1094. http://dx.doi.org/10.1176/appi.ps.52.8.1088

Jensen, P. S., Eaton Hoagwood, K., Roper, M., Arnold, L. E., Odbert, C., Crowe, M., . . . Wells, K. (2004). The services for children and adolescents-parent interview: Development and performance characteristics. *Journal of the American Academy of Child and Adolescent Psychiatry, 43,* 1334–1344. http://dx.doi.org/10.1097/01.chi.0000139557.16830.4e

Kazdin, A. E., & Blase, S. L. (2011). Rebooting psychotherapy research and practice to reduce the burden of mental illness. *Perspectives on Psychological Science, 6,* 21–37. http://dx.doi.org/10.1177/1745691610393527

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York, NY: Guilford Press.

Miranda, J., Bernal, G., Lau, A., Kohn, L., Hwang, W. C., & LaFromboise, T. (2005). State of the science on psychosocial interventions for ethnic minorities. *Annual Review of Clinical Psychology, 1,* 113–142. http://dx.doi.org/10.1146/annurev.clinpsy.1.102803.143822

Najavits, L. M., Gallop, R. J., & Weiss, R. D. (2006). Seeking safety therapy for adolescent girls with PTSD and substance use disorder: A randomized controlled trial. *The Journal of Behavioral Health Services & Research, 33,* 453–463. http://dx.doi.org/10.1007/s11414-006-9034-2

Nathan, P. E., & Gorman, J. M. (Eds.). (2007). *A guide to treatments that work* (3rd ed.). New York, NY: Oxford University Press.

Palinkas, L. A., Weisz, J. R., Chorpita, B. F., Levine, B., Garland, A. F., Hoagwood, K. E., & Landsverk, J. (2013). Continued use of evidence-based treatments after a randomized controlled effectiveness trial: A qualitative study. *Psychiatric Services, 64,* 1110–1118. http://dx.doi.org/10.1176/appi.ps.004682012

Park, A. L., Ebesutani, C. K., Bose, D., & Chorpita, B. F. (2015). Psychometric properties of a Spanish translation of the Revised Child Anxiety and Depression Scale—Parent version. *Journal of Psychopathology and Behavioral Assessment, 38,* 307–319. http://dx.doi.org/10.1007/s10862-015-9517-7

Park, A. L., Tsai, K. H., Guan, K., Reding, M. E. J., Chorpita, B. F., & Weisz, J. R., & the Research Network on Youth Mental Health. (2016). Service use findings from the Child STEPs effectiveness trial: Additional support for modular designs. *Administration and Policy in Mental Health and Mental Health Services Research, 43,* 135–140. http://dx.doi.org/10.1007/s10488-015-0625-1

Proctor, E. K., & Rosen, A. (2008). From knowledge production to implementation: Research challenges and imperatives. *Research on Social Work Practice, 18,* 285–291. http://dx.doi.org/10.1177/1049731507302263

Raghavan, R., Zima, B. T., Andersen, R. M., Leibowitz, A. A., Schuster, M. A., & Landsverk, J. (2005). Psychotropic medication use in a

national probability sample of children in the child welfare system. *Journal of Child and Adolescent Psychopharmacology, 15,* 97–106. http://dx.doi.org/10.1089/cap.2005.15.97

Reding, M. E. J., Guan, K., Regan, J., Palinkas, L. A., Lau, A. S., & Chorpita, B. F. (2016). *Implementation in a changing landscape: Provider experiences implementing evidence-based practice in Los Angeles County*. (Manuscript submitted for publication).

Roberts, R. E., Atrkisson, C. C., & Mendias, R. M. (1984). Assessing the Client Satisfaction Questionnaire in English and Spanish. *Hispanic Journal of Behavioral Sciences, 6,* 385–396. http://dx.doi.org/10.1177/07399863840064004

Rodríguez-Hernández, P. J., Betancort, M., Ramírez-Santana, G. M., García, R., Sanz-Álvarez, E. J., & De, C. (2012). Psychometric properties of the parent and teacher versions of the Strength and Difficulties Questionnaire (SDQ) in a Spanish sample. *International Journal of Clinical and Health Psychology, 12,* 265–279. Retrieved from http://search.proquest.com/docview/1285630430?accountid=14512

Rotheram-Borus, M. J., Swendeman, D., & Chorpita, B. F. (2012). Disruptive innovations for designing and diffusing evidence-based interventions. *American Psychologist, 67,* 463–476. http://dx.doi.org/10.1037/a0028180

Sanders, M. R., Markie-Dadds, C., Turner, K. M. T., & Ralph, A. (2004). *Using the Triple P system of intervention to prevent behavioural problems in children and adolescents. Handbook of interventions that work with children and adolescents: Prevention and treatment* (pp. 489–516). New York, NY: Wiley. http://dx.doi.org/10.1002/9780470753385.ch20

Schoenwald, S. K., & Hoagwood, K. (2001). Effectiveness, transportability, and dissemination of interventions: What matters when? *Psychiatric Services, 52,* 1190–1197. http://dx.doi.org/10.1176/appi.ps.52.9.1190

Southam-Gerow, M. A., Daleiden, E. L., Chorpita, B. F., Bae, C., Mitchell, C., Faye, M., & Alba, M. (2014). MAPping Los Angeles County: Taking an evidence-informed model of mental health care to scale. *Journal of Clinical Child and Adolescent Psychology, 43,* 190–200. http://dx.doi.org/10.1080/15374416.2013.833098

Stein, B. D., Jaycox, L. H., Kataoka, S. H., Wong, M., Tu, W., Elliott, M. N., & Fink, A. (2003). A mental health intervention for schoolchildren exposed to violence: A randomized controlled trial. *Journal of the American Medical Association, 290,* 603–611. http://dx.doi.org/10.1001/jama.290.5.603

Steinberg, A. M., Brymer, M. J., Decker, K. B., & Pynoos, R. S. (2004). The University of California at Los Angeles Post-traumatic Stress Disorder Reaction Index. *Current Psychiatry Reports, 6,* 96–100. http://dx.doi.org/10.1007/s11920-004-0048-2

Steinberg, A. M., Brymer, M. J., Kim, S., Briggs, E. C., Ippen, C. G., Ostrowski, S. A., . . . Pynoos, R. S. (2013). Psychometric properties of the UCLA PTSD reaction index: Part I. *Journal of Traumatic Stress, 26,* 1–9. http://dx.doi.org/10.1002/jts.21780

Stiffman, A. R., Horwitz, S. M., Hoagwood, K., Compton, W., III, Cottler, L., Bean, D. L., . . . Weisz, J. R. (2000). The Service Assessment for Children and Adolescents (SACA): Adult and child reports. *Journal of the American Academy of Child and Adolescent Psychiatry, 39,* 1032–1039. http://dx.doi.org/10.1097/00004583-200008000-00019

Tang, T. Z., & DeRubeis, R. J. (1999). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology, 67,* 894–904. http://dx.doi.org/10.1037/0022-006X.67.6.894

Taylor, S., & Todd, P. (1995). Assessing IT usage: The role of prior experience. *Management Information Systems Quarterly, 19,* 561–570. http://dx.doi.org/10.2307/249633

U.S. Department of Health and Human Services, Administration on Children, Youth and Families. (2012). *Promoting the safe, appropriate, and effective use of psychotropic medication for children in foster care*. [Memorandum]. Retrieved from http://www.acf.hhs.gov/sites/default/files/cb/im1203.pdf

Webster-Stratton, C. (2005). *The Incredible Years: A Trouble Shooting Guide for Parents of Children Aged 2-8*. Seattle, WA: The Incredible Years Inc.

Weisz, J. R., Chorpita, B. F., Frye, A., Ng, M. Y., Lau, N., Bearman, S. K., . . . the Research Network on Youth Mental Health. (2011). Youth top problems: Using idiographic, consumer-guided assessment to identify treatment needs and to track change during psychotherapy. *Journal of Consulting and Clinical Psychology, 79,* 369–380.

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,* 274–282. http://dx.doi.org/10.1001/archgenpsychiatry.2011.147

Weisz, J. R., Jensen-Doss, A., & Hawley, K. M. (2006). Evidence-based youth psychotherapies versus usual clinical care: A meta-analysis of direct comparisons. *American Psychologist, 61,* 671–689. http://dx.doi.org/10.1037/0003-066X.61.7.671