

The Large-Scale Implementation of Trauma-Focused Cognitive-Behavioral Therapy

(TF-CBT) in Community Practice Settings:

An Examination of Client and Implementation Outcomes

A Dissertation

SUBMITTED TO THE FACULTY OF THE

UNIVERSITY OF MINNESOTA

BY

Katelyn Marie Donisch

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Abigail Gewirtz, PhD

Bonnie Klimes-Dougan, PhD

July 2018

Katelyn M. Donisch, © 2018

Acknowledgements

This investigation was funded by the Substance Abuse and Mental Health Administration's National Child Traumatic Stress Network (awarded to Drs. Abi Gewirtz and Chris Bray) and the Minnesota Department of Human Services. Both entities supported the University of Minnesota's Ambit Network in establishing and sustaining trauma-informed practice and treatment across the state of Minnesota.

To the Ambit Network: I would like to recognize the exceptional efforts of the Ambit Network's administrative, training, and implementation team that made this investigation possible. Abi Gewirtz, your vision for the Ambit Network, and now the Center for Resilient Families, has led to considerable improvements in the delivery of trauma-informed care to youth, families, and communities not only in Minnesota, but throughout the United States. Your advising and guidance made it possible for me to pursue a graduate career in implementation and dissemination research. Chris Bray, you welcomed me into the Ambit Network six years ago, and supported and encouraged my love of program evaluation. I am so fortunate to call you a mentor and friend. Jennifer Wilgocki and Saprina Matheny, thank you for your unparalleled passion in training and coaching clinicians in Trauma-Focused Cognitive-Behavioral Therapy. To the Ambit Network's implementation team, Tanner Zimmerman and Katy Lund, your dedication made collecting, coding, and analyzing data for this investigation possible. I am also grateful to the clinicians who dedicated countless hours to learning and providing TF-CBT to their clients, as well as the clients who were courageous in discussing their trauma histories to begin the process of healing.

To my dissertation committee, Abi Gewirz, Ann Masten, Bonnie Klimes-Dougan, and Arthur Reynolds: You never questioned my desire to pursue implementation research, a field of study impacted by an inchoate research base and methodological complexities. Instead, you envisioned the benefit of this research to treatment developers, purveyors, and other implementation researchers.

To my parents and brother: I am incredibly grateful for your patience, love, and support. You have been present to assist me with every hardship, and celebrate every milestone. You all encouraged my curiosity from a young age by answering my myriad daily questions and cultivating my passion for inquiry. Your compassion towards, and professional devotion to, children and families continues to inspire me in my work with trauma-exposed children and adolescents.

To my husband: You first accompanied me to my graduate program in Minnesota, and then to my clinical internship in North Carolina. Thank you for supporting my personal and professional interests in every way possible—lengthy conversations, hikes in the woods, nightly meals, and delicious baked goods. Your patience, guidance, encouragement, and love made this dissertation possible. Thank you.

Dedication

This dissertation is dedicated to the children and adolescents with whom I have worked. Your resilience in the face of life's many obstacles is truly extraordinary.

Abstract

Although significant progress has been made in the development of trauma-focused evidence-based treatments (EBTs) for children and adolescents, the transfer of these interventions from academic to community practice settings remains challenging. This investigation is one of the first to assess the statewide implementation of a popular trauma-focused EBT for youth in routine practice settings. Using nine years of baseline and follow-up data, this investigation examined the implementation of Trauma-Focused Cognitive-Behavioral Therapy (TF-CBT) with 2,980 clients, 596 clinicians, and 75 community practice agencies. Clients were highly traumatized, reporting exposure to an average of four trauma types ($SD = 2.85$). Multilevel modeling indicated that clients participating in TF-CBT demonstrated statistically significant reductions in posttraumatic stress symptomatology from baseline to follow-up, as well as self- and parent-reported reductions in emotional symptoms, conduct problems, hyperactivity-inattention, peer problems, and the overall impact of these difficulties from baseline to follow-up. Clients also showed a statistically significant self- and parent-reported improvement in prosocial relations from baseline to follow-up. Although clinician TF-CBT implementation fidelity was not related to changes in client outcomes from baseline to follow-up, both clinician license and supervisory status were statistically significantly related to implementation fidelity score. Findings suggest that the large-scale implementation of TF-CBT is a promising approach to increasing the availability of trauma-focused EBTs in routine practice settings, while also reducing the primary and secondary effects of trauma in the lives of children and their families.

Table of Contents

List of Tables	vi
List of Figures	viii
Comprehensive Introduction	1
Study One	21
Introduction	21
Method	26
Results.....	38
Discussion	41
Tables	46
Figures	58
Study Two	59
Introduction	59
Method	66
Results	82
Discussion	90
Tables	98
Comprehensive Discussion	108
References	116

List of Tables

Study One

Table 1. Client Demographic Characteristics46

Table 2. UCLA PTSD-RI for DSM-IV and DSM-5 Item Comparisons:
Preparation for Common Item Equating47

Table 3. Intraclass Correlation Coefficients for Client Outcome Variables
at Baseline49

Table 4. Intraclass Correlation Coefficients for Client Outcome Variables
at Follow-Up50

Table 5. Mean and Total Number of Traumas Reported by Clients.....51

Table 6. Proportion of Sample Endorsing Specific Trauma Types52

Table 7. Client Outcomes at Baseline and Follow-Up53

Table 8. PTSD-RI: Estimated Fixed Effects and Covariance Parameters54

Table 9. SDQ Self-Report: Estimated Fixed Effects and Covariance
Parameters.....55

Table 10. SDQ Parent Report: Estimated Fixed Effects and Covariance
Parameters.....56

Study Two

Table 11. Agency Characteristics98

Table 12. Clinician Characteristics99

Table 13. Intraclass Correlation Coefficients for Client Outcome Deltas100

Table 14. Clinician Implementation Fidelity: Composite and Components
Score101

Table 15. Client Outcomes, Deltas, and Effect Sizes	102
Table 16. Bivariate Correlations (<i>r</i>) between Outcome Deltas and Implementation Fidelity Composite Scores.....	103
Table 17. PTSD-RI Deltas: Estimated Fixed Effects and Covariance Parameters.....	104
Table 18. SDQ Self-Report Deltas: Estimated Fixed Effects and Covariance Parameters.....	105
Table 19. SDQ Parent Report Deltas: Estimated Fixed Effects and Covariance Parameters.....	106
Table 20. Client, Clinician, and Agency Characteristics Associated with Trauma-Focused Cognitive-Behavioral Therapy (TF-CBT) Implementation Fidelity	107

List of Figures

Study One

Figure 1. Flow Diagram of Eligible Clients and Associated Baseline and Follow-Up Assessments.....	58
--	----

Comprehensive Introduction

National prevalence estimates suggest that approximately 60 percent of children and adolescents are exposed to child maltreatment or other traumatic events before the age of 17 (Finkelhor, Turner, Ormrod, & Hamby, 2009; Finkelhor, Turner, Shattuck, & Hamby, 2015). While many children and adolescents (hereafter referred to as youth) demonstrate resilience after trauma exposure (Cicchetti, Rogosch, Lynch, & Holt, 1993; Masten, 2001), other youth evince disruptions across developmental processes (Kaplow & Widom, 2007; Lansford et al., 2002). Specifically, research suggests that trauma-exposed children are more likely to manifest differences in neuroendocrine functioning (Cicchetti & Rogosch, 2001; Tarullo & Gunnar, 2006), attachment security (Carlson, Cicchetti, Barnett, & Braunwald, 1989; Pickreign Stronach, Toth, Rogosch, Oshri, Manly, & Cicchetti, 2011), neuropsychological functioning (i.e., attention, executive functioning; Beers & De Bellis, 2002), emotion recognition (Pollak, Cicchetti, Hornung, & Reed, 2000) and regulation (Kim & Cicchetti, 2010), among other atypicalities.

These differences pose an increased, if not significant, risk for the development of psychopathology across the life span, including posttraumatic stress disorder (PTSD; Deblinger, McLeer, Atkins, Ralphe, & Foa, 1989), anxiety (Johnsone et al., 2002), depression (Johnsone et al., 2002; Turner, Finkelhor, & Ormrod, 2006), substance abuse (Johnsone et al., 2002), and behavioral problems (Fairbank, Putnam, & Harris, 2007; Turner, Finkelhor, & Ormrod, 2006). Without effective treatment, symptoms of posttraumatic stress—as well as comorbid conditions—may persist over time (Hernandez, Lodico, & DiClemente, 1993; Hoven et al., 2005; Sigel & Williams, 2003). Moreover, research suggests that trauma-exposed youth and families who do not receive

adequate support may be more likely to experience increased rates of secondary adversities, such as physical health concerns (Widom, Czaja, Bentley, & Johnson, 2012), home and foster home placement disruptions (Aarons et al., 2010), academic difficulties (Shonk & Cicchetti, 2001), and social maladjustment (Shonk & Cicchetti, 2001).

Although substantial progress has been made to develop evidence-based treatments (EBTs) that address the primary and secondary consequences of PTSD in youth (Cohen, Deblinger, Mannarino, & Steer, 2004; Kolko & Swenson, 2002), program developers and purveyors have encountered difficulties in transporting EBTs to community practice settings where trauma-exposed youth are primarily served (Kolko, Cohen, Mannarino, Baumann, & Knudsen, 2009). In fact, among the estimated 15 million youth who meet diagnostic criteria for a mental disorder in the United States (American Psychological Association Task Force on Evidence-Based Practice for Children and Adolescents, 2008), only 20 (Kataoka, Zhang, & Wells, 2002) to 50 percent (Merikangas, He, Brody, Fisher, & Bourdon, 2010) receive mental health care each year, and even fewer children are reported to receive EBTs. These data, combined with the aforementioned high rates of trauma exposure for youth in the United States, suggest an urgent need to examine the implementation of trauma-focused EBTs in community practice settings.

Research on traditional and trauma-focused EBTs occurs on a continuum, generally progressing from investigations of efficacy and effectiveness to implementation efforts (Singal, Higgins, & Waljee, 2014). Specifically, efficacy research examines the performance of an intervention under ideal and controlled circumstances (e.g., homogenous client population, several exclusion criteria, highly trained providers),

whereas effectiveness research assesses an intervention's performance under "real world" practice conditions (e.g., heterogenous client population, few exclusion criteria, community mental health providers) (Schoenwald & Hoagwood, 2001; Singal et al., 2014). In comparison, implementation research is the study of the systematic uptake of EBTs into community, or routine, practice settings to improve overall quality of care (Bauer, Damschroder, Hagedorn, Smith, & Kilbourne, 2015). Therefore, as trauma-focused EBT developers and purveyors begin to transport their interventions into community practice settings, they must gradually progress from efficacy and effectiveness to implementation research.

To date, most investigations of trauma-focused EBTs for youth have evaluated efficacy and effectiveness rather than implementation in routine care settings. This investigation is one of the first to assess the large-scale implementation of a popular trauma-focused EBT for youth in diverse mental health care venues. Using nine years of baseline and follow-up data, this investigation examines the implementation of Trauma-Focused Cognitive-Behavioral Therapy in community practice settings, including outpatient, inpatient, residential treatment, and correctional agencies.

Trauma-Focused Cognitive-Behavioral Therapy (TF-CBT)

Trauma-Focused Cognitive-Behavioral Therapy (TF-CBT) is one of the most widely used EBTs for trauma-exposed youth in the United States (Cary & McMillen, 2014; Cohen, Mannarino, & Deblinger, 2017). TF-CBT is a components-based treatment model for youth ages 3 to 18 and their caregivers affected by trauma (Cohen, Mannarino, & Deblinger, 2006). Although TF-CBT was initially developed to address the sequelae of trauma associated with child sexual abuse (Deblinger, Mannarino, Cohen, & Steer, 2006),

the model has been adapted for use with youth exposed to a multiplicity of traumatic experiences: disasters (e.g., 2001 terrorist attack; CATS Consortium, 2007), maltreatment (Allen & Johnson, 2012), traumatic grief (Cohen, Mannarino, & Staron, 2006), chronic trauma (Cohen, Mannarino, & Murray, 2011), and complex trauma (Cohen, Mannarino, Kliethermes, & Murray, 2012). Other populations for which TF-CBT has been proven effective include young children (i.e., 3 to 6 years of age; Salloum, Scheeringa, Cohen, & Storch, 2014; Scheeringa, Weems, Cohen, Amaya-Jackson, & Guthrie, 2011), refugee youth (Murray, Cohen, Ellis, & Mannarino, 2008), children from low-income countries (Murray et al., 2013), and children with co-occurring trauma exposure and behavior problems (Cohen, Berliner, & Mannarino, 2010). In addition to treating PTSD, TF-CBT has been shown to improve co-occurring conditions, such as depression, anxiety, externalizing behavior problems, relationship and attachment problems, and academic difficulties (Cohen et al., 2006).

The TF-CBT model is comprised of eight essential elements: Psychoeducation and Parenting Skills (P), Relaxation (R), Affect Expression and Regulation (A), Cognitive Coping (C), Trauma Narrative Development and Processing (T), In Vivo Gradual Exposure (I), Conjoint Parent/Child Sessions (C), and Enhancing Safety/Future Development (E) (i.e., comprising a “PRACTICE” acronym). The PRACTICE components allow for a flexible, manualized treatment approach that can be divided into three phases of treatment (Cohen et al., 2017). Phase 1 consists of stabilization and skill building (i.e., “PRAC”), Phase 2 is characterized by trauma narration and processing (i.e., “T”), while Phase 3 is guided by consolidation and closure (i.e., “ICE”). Logistically, the TF-CBT treatment model consists of 12 to 25 treatment sessions, which are comprised of

individual youth and parent sessions, as well as conjoint youth-parent sessions (Cohen et al., 2006). Treatment sessions are generally 60 to 90 minutes in length and are conducted on a weekly basis at a pace determined by the youth's affective and cognitive readiness.

Across multiple treatment research reviews for youth with diagnoses of PTSD (i.e., California Evidence-Based Clearinghouse for Child Welfare, 2009; Chadwick Center for Children and Families, 2004; Saunders, Berliner, & Hanson, 2004; Silverman et al., 2008), TF-CBT was given the highest evidence-based rating (Sigel, Benton, Lynch, & Kramer, 2013). Indeed, TF-CBT was found to be effective in a large number of peer-reviewed studies—including 20 RCTs—in that children participating in TF-CBT treatment demonstrated improvements in PTSD, depression, and anxiety symptoms, with parents reporting reductions in parental distress and improvements in parenting skills (for reviews, see Cary & McMillen, 2012; Cohen et al., 2017; de Arellano et al., 2014). Similarly, in an analysis conducted by Silverman and colleagues (2008), TF-CBT was the only trauma-focused treatment found to be statistically significantly superior to psychosocial placebo or to another treatment in at least two group-design experiments, conducted in at least two independent research settings, by two independent investigatory teams. In summary, many empirical investigations, including RCTs, conclude that TF-CBT is an efficacious treatment for youth exposed to a variety of traumatic events, and has also been found to be effective in diverse clinical settings (Murray, Dorsey et al., 2013; Murray, Familiar et al., 2013).

Implementation of Trauma-Focused EBTs in Community Practice Settings

Although the development of EBTs for trauma-exposed youth and their families is a critical first step towards reducing the effects of trauma on children, families, and

communities (Fixsen, Blase, Duda, Naoom, & Van Dyke, 2010), the use of trauma-focused EBTs on a scale sufficient to benefit society requires additional empirical evidence documenting their effective implementation in settings where youth are primarily served (Kolko, Iselin, & Gully, 2011; McHugh & Barlow, 2010). Efforts to adopt trauma-focused EBTs are often carried out under the expectation that these interventions can be transferred to, or implemented in, community practice settings without modification. However, there exist vast differences between the contexts, conditions, and climates under which trauma-focused EBTs are developed (e.g., controlled clinical trials in university-based settings) and the characteristics of routine care settings where most mental health treatments are delivered (Ebert, Amaya-Jackson, Markiewicz, Kisiel, & Fairbank, 2012, p. 188; Fixsen et al., 2010; Proctor et al., 2009). Several conceptual frameworks have attempted to capture the multi-level, complex characteristics that support or impede EBT implementation across various settings (Damschroder et al., 2009; Durlak & DuPre, 2008; Proctor et al., 2009).

Proctor and colleagues (2009), for example, developed the Conceptual Model of Implementation Research (CMIR), which distinguishes between—and uniquely links—implementation processes and outcomes. Notably, CMIR distinguishes among intervention strategies, implementation strategies, and outcomes. Intervention strategies consist of EBTs that have been found to be efficacious and effective (e.g., TF-CBT), whereas implementation strategies are the specific efforts or activities that are designed to improve the uptake and sustainability of EBTs (Proctor et al, 2009). Finally, outcomes are the “effects of deliberate and purposive actions to implement new treatments, practices, and services,” and include client, service, and implementation outcomes

(Proctor et al., 2011, p. 65). Given CMIR's widespread use within the evaluation and implementation literature (Nilsen, 2015), it will be used to guide this investigation's examination of the large-scale implementation of TF-CBT through the assessment of both client and implementation outcomes.

Client outcomes. There exists a growing research literature describing the efficacy and effectiveness of TF-CBT with both outpatient and special populations of youth, including those in residential treatment (Cohen et al., 2016), foster care (Weiner, Schneider, & Lyons, 2009), and international settings (Murray, Dorsey et al., 2013; Murray, Familiar et al., 2013; Woods-Jaeger et al., 2017). However, very little TF-CBT research has examined the magnitude of client outcomes in the context of large-scale implementation efforts in community practice settings.

Lang and colleagues (Lang, Randall, Delaney, & Vanderploeg, 2017) recently published a descriptive examination of the Connecticut Department of Children and Families' nine-year initiative to implement TF-CBT in routine care settings. A total of 612 clinicians participated in TF-CBT training across 87 agencies, and served over 6,200 youth. Among youth who completed treatment, pre- and post-treatment analyses showed reductions in child PTSD symptoms as rated by both youth (44.8%) and their caregivers (41.0%); while for treatment non-completers, youth and their caregivers reported moderate PTSD symptom reductions (i.e., 25.0% by youth, 21.0% by caregivers). Reductions in child depression were also observed, as indicated by both the youth (47.0%) and caregivers (44.5%). Similarly, among youth who did not complete TF-CBT, youth and their caregivers reported moderate depressive symptom reductions (i.e., 24.9% by youth, 24.0% by caregivers). Although Lang and colleagues (2017) demonstrated that

TF-CBT implemented in routine practice settings resulted in reductions in youth- and caregiver-reported symptomatology, statistical analyses were only descriptive in nature, and did not account for length of treatment.

Implementation outcomes. As compared to client outcomes, implementation outcomes are defined as the “effects of deliberate and purposive actions to implement new treatments, practices, and services,” and consist of acceptability, adoption, appropriateness, cost (i.e., incremental or implementation cost), feasibility, fidelity, penetration, and sustainability (Proctor et al., 2011, p. 65). By studying various implementation outcomes, EBT developers and purveyors are better able to distinguish EBT effectiveness from implementation effectiveness. When implementation strategies fail, as they often do, it is important to ascertain whether the failure occurred because the intervention was ineffective in the new setting (i.e., intervention failure) or if an appropriate intervention was deployed incorrectly (i.e., implementation failure) (Proctor et al., 2011, p. 66). Therefore, the careful examination of implementation outcomes allows for the systematic transportation of EBTs from academic research settings to community-based mental health venues.

Implementation fidelity. Implementation fidelity is one of eight implementation outcomes described by Proctor et al. (2011) and is characterized as the degree to which a clinician delivers a treatment adherent to its core principles (Carroll et al., 2007; Proctor et al., 2011) or, in other words, the degree to which treatment models are delivered as intended by the program developers (Proctor et al., 2011). Implementation fidelity has been measured more often than any other implementation outcome given that insufficient fidelity often weakens intervention outcomes, leading to inaccurate conclusions about

intervention effectiveness (Proctor et al., 2011). Carroll and colleagues (2007) suggest the following about implementation fidelity:

It is only by making an appropriate evaluation of the fidelity with which an intervention has been implemented that a viable assessment can be made of its contribution to outcomes... Unless such an evaluation is made, it cannot be determined whether a lack of impact is due to poor implementation or inadequacies in the programme itself. (p. e1)

Distinguishing between implementation and intervention weaknesses is a resource- and time-intensive process (Fixsen et al., 2005); however, as EBT utilization continues to spread, understanding the level of implementation fidelity to specific trauma-focused EBT models is crucial for improving the care of trauma-exposed youth and their families

Notably, there exists a lack of consensus among implementation researchers as to what specifically constitutes implementation fidelity (Dane & Schneider, 1998). For instance, some experts proffer that implementation fidelity is composed of three components (i.e., treatment adherence, treatment differentiation, and treatment competence; Schoenwald et al., 2011), while others suggest up to eight (i.e., adherence, dosage, quality, participant responsiveness, program differentiation, monitoring of control/comparison conditions, program reach, and adaptation; Durlak & DuPre, 2008). As described in their conceptual framework for implementation fidelity, Carroll and colleagues (2007) identified four primary components of implementation fidelity, as well as four additional factors that may influence, or moderate, the degree of implementation fidelity.

According to Carroll et al.'s (2007) framework, the measurement of implementation fidelity is congruent with the measurement of *adherence* to an intervention, or the extent to which those responsible for delivering an intervention comply with the intervention as outlined by the model developers (p. 42). Adherence is comprised of four subcategories, including the content, frequency, duration, and coverage (i.e., dose) of an intervention. Specifically, an intervention's *content* is generally considered to be its core components, or the skills, knowledge, or elements of treatment that must be delivered to participants. The *frequency* of implementation is characterized as the intervention's rate of occurrence (e.g., daily, weekly, biweekly), while *duration* refers to length of intervention delivery (e.g., total number of sessions). Taken together, frequency and duration refer to the dosage of an intervention received by participants over time. Finally, *coverage* measures the extent to which "people who should be participating in or receiving the benefits of an intervention actually do so" (p. 41).

Of course, the degree to which an intervention is implemented with fidelity may be influenced by other variables. In their conceptual framework, Carroll et al. (2007) proffer that intervention complexity, facilitation strategies, quality of delivery, and participant responsiveness potentially moderate the relationship between implementation fidelity and intervention outcomes. However, in the implementation science literature, these variables may be considered core subcategories of implementation fidelity (see Dane & Schneider, 1998; Rohrbach, Dent, Skara, Sun, & Sussman, 2007). Given the demonstrated overlap between measures of adherence and quality of delivery, as well as the widespread recognition of Carroll et al.'s (2007) framework for implementation

fidelity, this dissertation defines implementation fidelity as consisting of the content, frequency, duration, and coverage of an intervention.

TF-CBT implementation fidelity. Within the TF-CBT implementation literature (i.e., excluding the efficacy literature), very few investigations have examined and reported on TF-CBT implementation fidelity (see Ebert et al., 2012; Lang et al., 2015; and Sigel et al., 2013 for examples). Largely, these investigations are focused on the feasibility of particular implementation strategies with little discussion of the relationship between implementation fidelity and client outcomes. As such, there is an urgent need for nuanced investigations of TF-CBT implementation fidelity in community practice settings.

Breakthrough Series Collaborative. Ebert et al. (2012) examined the feasibility and utility of adapting the Institute of Healthcare's Breakthrough Series Collaborative (BSC) to support the implementation and sustained use of TF-CBT in community practice settings. Among other evaluative components, a one-month implementation evaluation was conducted after the final TF-CBT learning session. TF-CBT implementation fidelity was measured in three ways: length of treatment, clinician skill, and adherence (i.e., content delivered). Length of treatment was assessed by requesting the number of TF-CBT sessions provided to clients that had completed treatment. In the BSC, the average length of treatment for completed cases was 17 sessions ($SD = 6.0$, $Median = 16$, $Range 7-32$ sessions). Clinician skill was measured using an 11-item TF-CBT Rating Form completed by supervisors for each supervisee. On the Rating Form, supervisors indicated, on a 5-point scale (1 = *never*, 5 = *almost always*), the frequency with which the supervisees implemented TF-CBT with a high degree of skill along

several key dimensions (e.g., implements treatment components in the order described in the TF-CBT manual unless there are compelling clinical reasons to alter). The overall mean was 4.2 ($SD = .61$), with mean scores for individual items ranging from a high of 4.5 ($SD = .63$; i.e., “Is able to identify clients for whom TF-CBT is an appropriate treatment”) to a low of 4.0 ($SD = .88$; “Is able to prioritize problems with families that have multiple problems so that TF-CBT remains the focus of the sessions”).

Finally, clinicians completed the TF-CBT Practice Checklist (Deblinger, Cohen, Mannarino, Hanson, & Runyon, 2005), a 49-item self-assessment of adherence to the TF-CBT treatment model. Specifically, the Practice Checklist assesses the frequency, on a 5-point scale (1 = *never*, 5 = *almost always*), with which the clinician engaged in specific activities used to deliver TF-CBT during the past two months (e.g., established an agenda and/or structure for the session, utilized behavioral rehearsal and/or role plays). Fifty-seven therapists from nine BSC teams completed the TF-CBT Practice Checklist. Subscale means ranged from a high of 3.8 ($SD = .43$) for therapy structure/style of session (e.g., utilized behavioral rehearsal and/or role plays) to a low of 3.0 ($SD = .94$) for personal safety skills (e.g., presented or reviewed names for private parts). The total scale mean was 3.6 ($SD = .57$), indicating relatively frequent use of all eight TF-CBT components, on average.

Connecticut Department of Children and Families. Lang and colleagues (Lang, Franks, Epstein, Stover, & Oliver, 2015) examined the first phase of TF-CBT implementation in Connecticut from 2007 to 2010. In order to study implementation fidelity, the investigators collected clinician-reported fidelity data on a monthly basis through the completion of a brief form of the TF-CBT Fidelity Checklist (Deblinger,

Cohen, Mannarino, Murray, & Epstein, 2007). The checklist identified which TF-CBT component was used during each session, and a total fidelity score was computed by computing the proportion of required treatment components as reported by clinicians over the entire duration of treatment. Lang and colleagues (2015) specified that conjoint witnessing and parenting sessions were not required for children who did not have a caregiver involved in treatment, and in-vivo exposure was not required for any children, as it is an optional clinical component. Fidelity data were available for 528 children (89.8% of the sample), but were only analyzed for the 231 children for whom treatment was deemed complete. Total fidelity scores for completed cases were quite high ($M = 0.936$, $SD = 0.119$), whereas reported completion rates for each component were as follows: Psychoeducation (94.8%), Parenting Skills Training (91.7% of cases with caregiver involvement), Relaxation Skills (96.5%), Affective Expression (97.4%), Cognitive Coping (96.5%), Trauma Narrative (98.3%), Conjoint Parent-Child Witnessing (89.5% of cases with caregiver involvement), and Enhancing Safety (84.8%). The optional In Vivo Exposure component was used in 53.7% of cases. In 2017, Lang and his team published updated implementation fidelity findings indicating that among clients completing TF-CBT, clinicians completed 89.0% of components, on average (Lang et al., 2017).

New England Residential Treatment Facilities. In a randomized implementation study of TF-CBT for adjudicated adolescents in residential treatment facilities, Cohen and colleagues (Cohen et al., 2016) required clinicians to complete a 9-item TF-CBT checklist immediately after each session to indicate which TF-CBT component was delivered during the session. Clinicians were randomized to participate in one of two TF-

CBT implementation strategies: 1) web-based TF-CBT training with an online consultation course (W), or 2) the W strategy in addition to a two-day in-person TF-CBT workshop, and 12 months of twice-monthly TF-CBT expert consultation calls with behavioral rehearsal (W+L). Regarding rates of intervention completion, results indicated that significantly more adjudicated youth receiving TF-CBT from W+L therapists completed treatment compared with youth treated by W therapists. Further, a significantly higher proportion of youth treated by W+L therapists received TF-CBT with high fidelity ratings than those treated by W therapists. However, among adjudicated youth who *completed* treatment, the investigators found no significant differences in implementation fidelity between W and W+L cases. Regardless of implementation strategy, results from the investigation indicated that youth who completed TF-CBT experienced statistically and clinically significant reductions in PTSD- and depression-related symptomatology from pre- to post-treatment.

Relationship between client outcomes and implementation fidelity.

Researchers across disciplines have begun to investigate the impact of implementation fidelity on client outcomes. For example, a 2008 review of five meta-analyses investigated the impact of various prevention and intervention programs (e.g., anti-bullying programs, drug use prevention studies, school-based prevention programs targeting aggressive behaviors) on client outcomes (Durlak & DuPre, 2008). Results revealed that interventions with better, as opposed to poorer, implementation demonstrated effect sizes that were two to three times larger. Under ideal circumstances, these effect sizes increased to twelve times larger with better implementation fidelity (Derzon, Sale, Springer, & Brounstein, 2005). Of course, such review and meta-analytic

findings are limited by the data presented in the original investigations, particularly with respect to the precise level of implementation achieved (Durlak & DuPre, 2008). To date, no large-scale TF-CBT implementation study has directly investigated the relationship between implementation fidelity and client outcomes in community practice settings. In the current study, this gap is addressed by examining the association between TF-CBT implementation fidelity and a variety of client outcomes in a statewide sample.

Factors Affecting Implementation Fidelity

In addition to a burgeoning interest in the assessment of implementation fidelity, a large body of work has investigated the core components, or essential elements, of the successful implementation of EBTs (Cohen et al., 2016; Fixsen, Naoom, Blasé, Friedman, & Wallace, 2005). However, a variety of factors affect the implementation of EBTs in community practice settings, and these factors occur across multiple levels of the practice context, including the client (Fixsen et al., 2005), clinician (Durlak & DuPre, 2008; Fixsen et al., 2005; Kolko et al., 2009; Ruzek & Rosen, 2009), and organizational levels (Aarons, Sommerfeld, & Walrath-Greene, 2009; Fixsen et al., 2005). Few TF-CBT researchers have investigated the relationship between implementation fidelity and client outcomes; more have studied the clinician-level factors that are associated with implementation fidelity (see Allen & Johnson, 2008; Cohen et al., 2016; Fritz et al., 2013; Lang et al., 2015; Woody, Anderson, & D'Souza, 2015).

Client-level factors. Research is only beginning to systematically evaluate the effects of various client demographic and trauma characteristics that may moderate the effectiveness of trauma-focused EBTs. Even fewer investigations study the effect of client characteristics on trauma-focused EBT and TF-CBT implementation fidelity. For

example, results from a study examining the implementation of Cognitive Therapy for PTSD in adults suggest that clients treated for multiple traumas were more likely to attend irregularly compared with clients who were treated for one or two traumas (Ehlers et al., 2013). Ehlers et al. (2013) also discovered that the treatment for clients with multiple traumas was less trauma-focused than the treatment of other clients, although the authors did not find statistical differences in dropout rate or total number of treatment sessions completed.

Clinician-level factors. Significant variability exists in the extent to which clinician-level characteristics (e.g., demographic characteristics, skills competence) affect implementation fidelity. For example, in a survey of 132 mental health clinicians practicing at children's advocacy centers, clinicians' professional discipline (i.e., counseling, social work, clinical psychology) was not related to the likelihood of regularly implementing all TF-CBT components (Allen & Johnson, 2012). This result was replicated by Cohen and colleagues (2016) in that professional discipline did not significantly impact TF-CBT implementation outcomes. Similarly, no discipline differences (e.g., degree, type of license) were found to exist between participants and nonparticipants in TF-CBT consultation calls (i.e., as studied in AR BEST's model; Fritz et al., 2013). Regarding licensure status, licensed therapists have been found to be significantly more likely than non-licensed therapists to assess for PTSD and depression symptoms, engage and retain youth in TF-CBT treatment, and to complete TF-CBT with fidelity, with a trend for licensure status to also impact completion of TF-CBT (Cohen et al., 2016).

Regarding additional demographic predictors, Allen and Johnson (2012) discovered that clinician age was not associated with the likelihood of regularly implementing all TF-CBT components. In contrast, another study found that TF-CBT knowledge scores (i.e., assessed by correct responses to hypothetical clinical vignettes) were statistically significantly positively related to clinician age (Woody, Anderson, D'Souza, Baxter, & Schubauer, 2015). Interestingly, clinician age was also statistically significantly related to referred clients coming to their first scheduled appointment (Woody, Anderson, D'Souza, 2015). It was posited that older, more experienced practitioners may have received more referrals, which allowed for more opportunities to apply, and gain knowledge about, the TF-CBT model.

Other studies have investigated the relationship between clinicians' total years of experience and TF-CBT implementation. For instance, across multiple investigations, clinicians' years of experience did not predict the likelihood of implementing all TF-CBT components (Allen & Johnson, 2012) and/or implementation outcomes (Cohen et al., 2016), and were not associated with participation in TF-CBT consultation calls during the AR BEST initiative (i.e., as predicted by number of years licensed; Fritz et al., 2013). In contrast, in a separate investigation, clinician experience was statistically significantly positively associated with recent self-ratings of competence in TF-CBT (i.e., global rating of competence in most recent TF-CBT session; Woody, Anderson, D'Souza, 2015). Also, clinician experience was positively associated with post-training self-efficacy scores in therapy involving parents, and therapy involving joint parent-child trauma work (Woody, Anderson, D'Souza, 2015a).

Organization-level factors. Although organization-level characteristics have not been included in TF-CBT implementation fidelity research, studies in the substance abuse field indicate that several organization-level factors contribute to clinicians' poorer implementation fidelity when implementing evidence-based substance abuse treatments (e.g., leadership attitudes, staff resources, organizational climate, management style), even with the provision of individualized consultation (for a review, see Simpson, 2002). Further, Glisson and Hemmelgarn (1998) have demonstrated that organizational climate—the psychological impact of the work environment—affects the outcomes of intensive case management services for children in the child welfare system (Glisson & Hemmelgarn, 1998) and the implementation of family-centered care in hospital emergency rooms (Hemmelgarn, Glisson, & Dukes, 2001).

Current Investigation

With the continued utilization of TF-CBT in national and international communities, it is important to document effective implementation strategies for transporting the TF-CBT model into accessible community practice settings, while simultaneously investigating associated implementation research outcomes (i.e., implementation and client outcomes). This research is integral to support the competent implementation and sustainability of TF-CBT on a scale sufficient to not only benefit trauma-exposed children and families, but also communities and larger society.

The present investigation adds to the TF-CBT literature by examining the multi-year, statewide implementation of TF-CBT by the University of Minnesota's Ambit Network, a grantee of the congressionally mandated National Child Traumatic Stress Network (NCTSN). Using a non-experimental prospective longitudinal design, this

investigation presents data from 2,980 trauma-exposed clients served by 596 clinicians across 75 agencies. Given its size, this is one of the largest initiatives to assess both TF-CBT client and implementation outcomes in the context of routine mental health practice, and was designed to address three research questions.

First, is TF-CBT associated with a decrease in posttraumatic stress, emotional problems, and behavioral problems (i.e., conduct problems, hyperactivity-inattention, peer problems) from baseline to follow-up in a community sample of youth? This question seeks to replicate and extend previous findings from the TF-CBT implementation literature with a more nuanced examination of the complex, non-independent data inherent to implementation (i.e., timepoints nested within clients, clients nested within clinicians, clinicians nested within agencies). Thus, this investigation will be the first to utilize multilevel modeling to account for potential clustering of the data. This question is addressed in Study 1: “The Large-Scale Implementation of Trauma-Focused Cognitive-Behavioral Therapy in Community Practice Settings: An Examination of Client Outcomes.”

The second and third questions address TF-CBT implementation outcomes. Specifically, is clinician implementation fidelity associated with posttraumatic stress, emotional problem, and behavioral problem (i.e., conduct problems, hyperactivity-inattention, peer problems) change scores from baseline to follow-up in a community sample of youth? And, third, what client-, clinician-, and organization-level factors are associated with implementation fidelity to the TF-CBT treatment model? The second and third questions are explored in Study 2: “The Large-Scale Implementation of Trauma-

Focused Cognitive-Behavioral Therapy (TF-CBT) in Community Practice Settings: An Assessment of Implementation Fidelity.”

The Large-Scale Implementation of Trauma-Focused Cognitive-Behavioral Therapy (TF-CBT) in Community Practice Settings: An Examination of Client Outcomes

Although substantial progress has been made to develop evidence-based treatments (EBTs) that address the primary and secondary consequences of posttraumatic stress in children and adolescents (Cohen et al., 2004; Kolko & Swenson, 2002), program developers and purveyors have encountered difficulties transporting these EBTs to community practice settings where trauma-exposed youth are primarily served (Kolko et al., 2009). In fact, among the estimated 15 million youth who meet diagnostic criteria for a mental disorder in the United States (American Psychological Association Task Force on Evidence-Based Practice for Children and Adolescents, 2008), only 20 (Kataoka et al., 2002) to 50 percent (Merikangas et al., 2010) receive mental health care each year and even fewer youth are reported to receive EBTs. Among youth with serious emotional disturbance, penetration rates of EBTs are estimated to be only one to three percent (Bruns et al., 2015). Therefore, while the development of EBTs for trauma-exposed youth and their families is a critical first step towards reducing the effects of trauma on children, families, and communities (Fixsen et al., 2010), the use of trauma-focused EBTs on a scale sufficient to benefit society requires additional empirical evidence documenting their effective implementation in community practice settings (Kolko et al., 2011; McHugh & Barlow, 2010).

Efforts to adopt trauma-focused EBTs are often carried out under the expectation that these interventions can be transferred to, or implemented in, community practice settings without adaption or tailoring. In efficacy trials, for instance, study procedures are

designed to streamline client presentations (i.e., exclude clients with comorbidities), enhance clinician skill, and reduce clinical demands; as well as increase resources within the service setting to maximize outcomes (Chambers et al., 2013). There exist vast differences, however, between the contexts, conditions, and climates under which trauma-focused EBTs are developed (e.g., controlled clinical trials in university-based settings) and the characteristics of routine care settings where mental health treatments are delivered (Ebert et al., 2012, p. 188; Fixsen et al., 2010; Proctor et al., 2009).

For example, research has found that youth referred to community practice settings evidence higher levels of co-occurring problems and less familial support than samples from university-based research clinics (Ehrenreich-May et al., 2011; Schoenwald & Hoagwood, 2001; Shirk, Karver, & Brown, 2001; Southam-Gerow, Weisz, & Kendall, 2003). Moreover, clinicians in community practice settings tend to provide treatment for a range of disorders with larger caseloads (Kazdin, 2002; Southam-Gerow, Rodriguez, Chorpita, & Daleiden, 2012; Weisz & Addis, 2006; Weisz & Gray, 2008). Given these differences, many interventions fail to achieve the same magnitude of effect observed in efficacy and effectiveness trials once they are implemented in the community. This phenomenon is referred to as “voltage drop” (Chambers et al., 2013; Kilbourne et al., 2007).

Trauma-Focused Cognitive-Behavioral Therapy (TF-CBT)

Trauma-Focused Cognitive-Behavioral Therapy (TF-CBT) is one of the most widely used EBTs for trauma-exposed youth in the United States (Cary & McMillen, 2014; Cohen et al., 2017). TF-CBT is a components-based treatment model for youth ages 3 to 18 and their caregivers affected by trauma (Cohen et al., 2006). Although TF-

CBT was initially developed to address the sequelae of trauma associated with child sexual abuse (Deblinger et al., 2006), the model has been adapted for use with youth exposed to a multiplicity of traumatic experiences: disasters (e.g., 2001 terrorist attack; CATS Consortium, 2007), maltreatment (Allen & Johnson, 2012), traumatic grief (Cohen et al., 2006), chronic trauma (Cohen et al., 2011), and complex trauma (Cohen et al., 2012). Other populations for which TF-CBT has been proven effective include young children (i.e., 3 to 6 years of age; Salloum et al., 2014; Scheeringa et al., 2011), refugee youth (Murray et al., 2008), children from low-income countries (Murray, Dorsey et al., 2013), and children with co-occurring trauma exposure and behavior problems (Cohen et al., 2010). In addition to treating PTSD, TF-CBT has been shown to improve co-occurring conditions, such as depression, anxiety, externalizing behavior problems, relationship and attachment problems, and academic difficulties (Cohen et al., 2006).

The TF-CBT model is comprised of eight essential elements: Psychoeducation and Parenting Skills (P), Relaxation (R), Affect Expression and Regulation (A), Cognitive Coping (C), Trauma Narrative Development and Processing (T), In Vivo Gradual Exposure (I), Conjoint Parent/Child Sessions (C), and Enhancing Safety/Future Development (E). Summarized by the acronym PRACTICE, TF-CBT's components allow for a flexible, manualized treatment approach that can be divided into three phases of treatment (Cohen et al., 2017). Phase 1 (i.e., PRAC) consists of stabilization through psychoeducation about trauma and traumatic stress, as well as skills training in relaxation, emotion regulation, and cognitive coping. Concurrent parent sessions in this phase assist in the functional analysis of youth behavior, as well as the introduction of selective attention, time-out procedures, and contingency reinforcement programs (Cohen

et al., 2017). Phase 2 (i.e., T) is characterized by the development of a trauma narrative and cognitive processing of the youth's trauma-related maladaptive cognitions, while Phase 3 (i.e., ICE) provides structure for consolidation and closure through psychoeducation review, sharing of the trauma narrative, and enhancement of safety (Cohen et al., 2017). The treatment model consists of 12 to 25 weekly sessions, which are generally 60 to 90 minutes in length and are conducted on a weekly basis at a pace determined by the youth's affective and cognitive readiness (Cohen et al., 2006).

Across multiple treatment research reviews for youth with diagnoses of PTSD (i.e., California Evidence-Based Clearinghouse for Child Welfare, 2009; Chadwick Center for Children and Families, 2004; Saunders et al., 2004; Silverman et al., 2008), TF-CBT was given the highest evidence-based rating (Sigel et al., 2013). Indeed, TF-CBT was found to be effective in a large number of peer-reviewed studies—including 20 RCTs—in that children participating in treatment demonstrated improvements in PTSD, depression, and anxiety symptoms, with parents reporting reductions in parental distress and improvements in parenting skills (for reviews, see Cary & McMillen, 2012; Cohen et al., 2017). With respect to effect sizes, medium to large effects have been reported at posttreatment with improvements maintained for up to two years (Deblinger et al., 2006; Deblinger, Steer, & Lippmann, 1999). In an analysis conducted by Silverman and colleagues (2008), TF-CBT was the only trauma-focused treatment found to be statistically significantly superior to psychosocial placebo or to another treatment in at least two group-design experiments, conducted in at least two independent research settings, by two independent investigatory teams. In summary, many empirical investigations, including RCTs, conclude that TF-CBT is an efficacious and effective

treatment for youth exposed to a variety of traumatic events, and has also been found to be efficacious in diverse clinical settings (Murray, Dorsey et al., 2013; Murray, Familiar et al., 2013).

Effectiveness and Implementation of Trauma-Focused Cognitive-Behavioral Therapy

Given TF-CBT's documented efficacy and effectiveness in a variety of community practice settings (i.e., residential treatment settings [Cohen et al., 2016], foster care [Weiner et al., 2009]), its widescale implementation represents one possible approach to improving the treatment of trauma-exposed youth and their families at a societal level (Sigel et al., 2013). However, very little TF-CBT research has examined youth outcomes in the context of large-scale implementation efforts in community practice settings.

Recently, Lang and colleagues (Lang et al., 2017) published a descriptive examination of the Connecticut Department of Children and Families' nine-year initiative to implement TF-CBT in routine care settings. A total of 612 clinicians participated in TF-CBT training across 87 agencies, and served over 6,200 youth. Among youth who completed treatment, pre- and post-treatment analyses showed reductions in child PTSD symptoms as rated by both youth (44.8%) and their caregivers (41.0%); while for treatment non-completers, youth and their caregivers reported moderate PTSD symptom reductions (i.e., 25.0% by youth, 21.0% by caregivers). Reductions in child depression were also observed, as indicated by both the youth (47.0%) and caregivers (44.5%). Again, among youth who did not complete TF-CBT, youth and their caregivers reported moderate depressive symptom reductions (i.e., 24.9% by youth, 24.0% by caregivers).

Although Lang and colleagues (2017) demonstrated that the implementation of TF-CBT in routine practice settings was associated with reductions in youth- and caregiver-reported symptomatology, statistical analyses were only descriptive in nature, and did not account for length of treatment.

Current Study

This study adds to the TF-CBT implementation literature by examining the multi-year statewide implementation of TF-CBT in community practice settings by the University of Minnesota's Ambit Network. Using a non-experimental prospective research design, this study presents baseline and follow-up data from 2,980 trauma-exposed clients served by 596 clinicians across 75 agencies. Given its size, this study is one of the largest to assess TF-CBT client outcomes in the context of routine mental health practice. To gain a better understanding of the typical child served in community practice settings in Minnesota, the author will first examine the number and types of traumas experienced by clients in routine care settings. Second, the author will determine whether client participation in TF-CBT is associated with a decrease in posttraumatic stress, emotional problems, and behavioral problems (i.e., conduct problems, hyperactivity-inattention, peer problems) from baseline to follow-up. Based on previous TF-CBT research in community practice settings (Lang et al., 2017), it is hypothesized that participation in TF-CBT will be associated with a decrease in posttraumatic stress, emotional problems, and behavioral problems from baseline to follow-up.

Method

Background

The University of Minnesota's Ambit Network was a National Child Traumatic Stress Network Category III Community Treatment and Services Center from 2005 to 2016. In partnership with the Minnesota Department of Human Services (DHS), the Ambit Network worked to establish and sustain trauma-informed practice and treatment across diverse client populations (e.g., military, refugee, and American Indian children) and service systems (e.g., mental health, child welfare, juvenile justice, and education) through targeted workforce development. From 2007 to 2016, the Ambit Network implemented over 21 learning collaboratives to train mental health clinicians in TF-CBT. The study protocol was approved by the University of Minnesota's Institutional Review Board and in order to maintain confidentiality, assessment measures were deidentified by clinicians before they were processed by the Ambit Network. Fidelity measures did not include any type of private or confidential client data.

Participants

Clients. TF-CBT was offered to a total of 3,632 clients over the course of 21 learning collaboratives. Although clients as young as 3 years of age participated in TF-CBT, the study sample was limited to clients between the ages of 7 and 18 given age cutoffs for the UCLA PTSD-Reaction Index for the *Diagnostic and Statistical Manual for Mental Disorders, 4th Edition (DSM-IV)*; American Psychiatric Association, 2000) and *5th Edition (DSM-5)*; American Psychiatric Association, 2013). Specifically, 305 clients were excluded from analyses as they were younger than age 7, and 12 clients were excluded as they were older than age 18. Analyses also excluded 212 clients who had participated in an early TF-CBT learning collaborative (i.e., clients were only administered the UCLA PTSD-Reaction Index for *DSM-IV* without the collection of

associated demographic and fidelity data), as well as 123 clients with multiple case openings. For clients with multiple case openings, the author used a decision tree to select the most recent and/or complete case opening. Using the aforementioned criteria, 2,980 clients between the ages of 7 and 18 completed an assessment to determine the appropriateness of TF-CBT. In accordance with the Strengthening the Reporting of Observational Studies in Epidemiology Statement (von Elm et al., 2007), refer to Figure 1 for a flow diagram of eligible clients and associated baseline and follow-up assessments.

Clients presented for treatment given a range of concerns and were identified by the clinicians as appropriate for TF-CBT during screening and assessment. Client demographics are presented in Table 1. Clients were an average of 13.0 years of age ($SD = 3.1$) and primarily identified as female (56.5%; $n = 1,684$). The racial and ethnic makeup of the clients was diverse. The majority of clients identified as Caucasian (60.5%; $n = 1804$), while 11.9% identified as Black or African American ($n = 355$), 8.1% as American Indian or Alaska Native ($n = 241$), 1.3% as Asian ($n = 40$), 0.4% as Native Hawaiian or Pacific Islander ($n = 12$), and 9.7% as multiracial ($n = 288$). Approximately 10% of youth identified as Hispanic ($n = 282$, 9.5%). Clients primarily resided with at least one biological or adoptive parent (55.8%; $n = 1,664$), while 16.1% lived in a residential treatment center ($n = 480$), 10.5% in kinship care ($n = 312$), 9.5% in foster care (i.e., traditional or treatment; $n = 284$), and 0.8% in a correctional facility ($n = 23$). Approximately 7.2% of clients reported that they resided in another setting (e.g., on their own, homeless, shelter; $n = 216$).

Clinicians and agencies. A total of 596 clinicians across 75 agencies provided TF-CBT to clients in community practice settings. Most agencies offered both outpatient and in-home mental health services ($n = 33$, 44.0%); followed by outpatient only ($n = 13$, 17.3%); outpatient, in-home, and inpatient ($n = 11$, 14.7%); outpatient and inpatient ($n = 7$, 9.3%); inpatient only ($n = 5$, 6.7%); comprehensive (i.e., outpatient, in-home, inpatient, and correction/residential treatment; $n = 3$, 4.0%); correction/residential treatment only ($n = 2$, 2.7%); and in-home only ($n = 1$, 1.3%) services. For the subsample of clinician-client pairs with fidelity data ($n = 629$, 21.1%), treatment lasted an average of 16.2 sessions ($SD = 11.8$; Range = 1-71) before completion or termination (i.e., client moved, client lost to follow-up, etc.).

Measures

Enrollment questionnaire. Demographic data included client age, gender (Male, Female, Transgender Male/Transgender Female, Other), race (American Indian or Alaska Native, Asian, Black/African American, Native Hawaiian or Other Pacific Islander, White, Unknown), ethnicity (Hispanic or Latino, Not Hispanic or Latino, Unknown), and current primary residence (home with parent(s), residential treatment center, kinship care [with relatives or other family], foster care [treatment and traditional foster care], correctional facility, other [independent, homeless, shelter, other]).

Assessments of client symptomatology, functioning, and trauma exposure. Data were collected using two to four psychometrically valid assessments per client depending on client age. Assessment measures were administered as face-to-face interviews by clinicians, although the *Strengths and Difficulties Questionnaire* could be administered via paper and pencil reporting. Assessments were administered and collected at baseline

(i.e., pre-treatment), every three months, and immediately at post-treatment. For the purpose of this study and to remain consistent with the TF-CBT implementation literature, baseline and closing assessments were utilized; if a closing assessment was not completed, the last follow-up assessment was carried forward as a closing assessment (see Lang et al., 2015 for similar procedures). Clinicians sent the completed assessments to Ambit Network staff via fax or email.

North Shore Trauma History Checklist (NTHC). The NTHC is a 30-item questionnaire that assesses for 17 types of trauma exposure through “Yes/No” responses to questions (North Shore-Long Island Jewish Health System, Inc., 2006). Respondents are also asked about the age of trauma onset, as well as the duration of the trauma exposure. The NTHC is completed by the interviewing clinician, and includes reports from both the child, caregiver, and/or other collateral sources. Trauma types include: disaster, terrorism/political violence, war, serious accident injury, illness/medical trauma, physical assault, physical maltreatment/abuse, neglect, separation from caregiver, homelessness, domestic violence, extreme interpersonal violence, bereavement, community violence, kidnapping/abduction, sexual maltreatment/abuse, sexual assault/rape, and other trauma not reported elsewhere. The last question asks the respondent to determine the “most significant traumatic event(s)” with the following prompt: “Of the things we’ve talked about, which is the worst? Which still really bothers you?”

General Trauma Information Form (GTF). The GTF assesses 22 different types of trauma exposure, in addition to salient characteristics of each exposure (e.g., age, frequency, confirmed or suspected exposure; National Child Traumatic Stress Network

2014). The GTF is completed by the interviewing clinician, and includes reports from both the child, caregiver, and/or other collateral sources. Trauma types include: (1) Sexual maltreatment/abuse: actual or attempted sexual molestation, exploitation or coercion by a caregiver; (2) Sexual assault/rape: actual or attempted sexual molestation, or coercion not by a caregiver and not recorded as sexual abuse; (3) Trafficking: child has been offered, obtained, or engaged in any kind of trafficking in exchange for anything of value, money, goods, personal benefit, in-kind favors, or some other kind of benefit; (4) Physical maltreatment/abuse: actual or attempted infliction of physical pain or bodily injury by a caregiver; (5) Physical assault: actual or attempted infliction of physical pain or bodily injury not by a caregiver and not recorded as physical abuse; (6) Psychological or emotional maltreatment/abuse: verbal abuse, excessive demands, emotional neglect; (7) Neglect: physical, medical, or educational neglect; (8) Domestic violence: exposure to physical, sexual, and/or emotional abuse directed at adult caregiver(s) in the home; (9) War/terrorism/political violence inside the U.S.: exposure to any of these events inside the U.S.; (10) War/terrorism/political violence outside the U.S.: exposure to any of these events outside the U.S.; (11) Illness/medical trauma: chronic, life-threatening, or extremely painful illness or medical procedure; (12) Serious accident/injury: unintentional accident or injury; (13) Disaster: major accident or disaster that is the result of a natural or manmade event; (14) Kidnapping/abduction: unlawful seizure or detention against the child's will; (15) Bereavement: death of a loved one, primary caregiver, or sibling; the unexpected, or premature death of a close relative or close friend; (16) Separation: separation from a primary caregiver or sibling; (17) Forced displacement: forced relocation to a different country due to political reasons; (18) Impaired caregiver:

history of exposure to caretaker depression/psychiatric problems, other medical illness, or alcohol/drug abuse; (19) Extreme interpersonal violence not reported elsewhere: for example, homicide or suicide; (20) Community violence not reported elsewhere: for example, gang-related violence, neighborhood violence; (21) School violence not reported elsewhere: for example, school shooting, classmate suicide; (22) Bullying: unwanted aggressive behavior among school aged children/peers that involves a real or perceived power imbalance; (23) Other trauma not reported elsewhere. Additionally, the clinician is asked to specify the primary focus of the current treatment using the 23 aforementioned traumas.

UCLA Posttraumatic Stress Disorder—Reaction Index for DSM-IV (UCLA PTSD-RI-IV). The UCLA PTSD-RI-IV is a 48-item questionnaire that assesses for trauma exposure and symptoms of PTSD in children and adolescents, and demonstrates robust reliability and validity (see Elhai et al., 2013; Pynoos, Rodriguez, Steinberg, Stuber, & Frederick, 1998; Steinberg et al., 2004; Steinberg et al., 2013). In this sample, Cronbach’s alpha was .619 at baseline. Responses are on a 5-point Likert scale, with higher ratings indicating more frequent PTSD symptoms (i.e., 0 = *none of the time*, 4 = *most of the time*). Only the 22-item symptom scales were used in this study. Specifically, 20 items assess PTSD symptoms, while two additional items screen for associated features: fear of recurrence, and trauma-related guilt. Overall, the UCLA PTSD-RI-IV provides an overall PTSD severity score, and separate scores for each PTSD criterion: Criterion B (Re-Experiencing), Criterion C (Avoidance/Numbing), and Criterion D (Arousal).

UCLA Posttraumatic Stress Disorder—Reaction Index for DSM-5 (UCLA PTSD-RI-5). The UCLA PTSD-RI-5 is a 31-item questionnaire that assesses for symptoms of

PTSD in children and adolescents (Pynoos & Steinberg, 2013). Although it includes sections on trauma history and exposure details, only the symptom scales were used in this study. Similar to the UCLA PTSD-RI-IV, responses are on a 5-point Likert scale, with higher ratings indicating more frequent symptoms (i.e., 0 = *none of the time*, 4 = *most of the time*). Specifically, 27 items assess PTSD symptoms, while four additional items screen for the dissociative subtype. Overall, the UCLA PTSD-RI-5 provides an overall severity score, and separate scores for each PTSD symptom category: Category B (Intrusion Symptoms), Category C (Avoidance of Stimuli), Category D (Negative Alterations in Cognitions and Mood), and Category E (Marked Alterations in Arousal and Reactivity). Reliability and validity have not yet been published for this measure; however, in this sample, Cronbach's alpha was .770 at baseline.

Strengths and Difficulties Questionnaire—Parent and Self-Report Versions (SDQ). The SDQ is a brief, 25-question behavioral screening measure for children and adolescents with good reliability (Goodman, 2001). Three versions of the SDQ were made available to clinicians, including the Parent and Teacher Versions, which were used with children 3 to 16 years of age, and the Self-Report Version, which was used with children 11 to 16 years of age. Each version has five clinical scales with five items each: Emotional Symptoms, Conduct Problems, Hyperactivity/Inattention, Peer Relationship Problems, Prosocial Behavior, and Impact. Each item is rated on a 3-point Likert scale with higher ratings indicating the extent to which each attribute applies to the child (i.e., 0 = *not true*, 2 = *certainly true*).

Data Analysis

Analyses were conducted using SPSS 24.0 (IBM Corp., 2016), including descriptive analyses, common item equating, and multilevel modeling. Distribution of dependent variable normality was checked using measures of skewness, kurtosis, and histogram assessment. Skewness and kurtosis values were within ± 2 and histograms showed normal distributions for all outcome variables studied (Curran, West, & Finch, 1995). SDQ Self-Report Prosocial Relations showed negative skewness ($S = -.855$, $SE = .171$), while PTSD-RI follow-up ($S = 0.866$, $SE = .100$) showed positive skewness, so data were transformed and re-analyzed. The results were similar to those using raw data; therefore, in this study, results using raw data are presented. Effect sizes were calculated using an online effect size calculator (Ellis, 2009) and interpreted using Cohen's (1988) original magnitude estimation for Cohen's d : $0.20 = small$, $0.50 = medium$, $0.80 = large$.

Common item equating. In 2013, the American Psychiatric Association (APA) published new diagnostic criteria for PTSD (APA, 2013) resulting in the 2015 replacement of the UCLA PTSD-RI-IV (Pynoos et al., 1998) with the UCLA PTSD-RI-5 (Pynoos & Steinberg, 2013). In line with this change, the Ambit Network adopted the UCLA PTSD-RI-5 in 2015 after using the UCLA PTSD-RI-IV from 2007 to 2015. Approximately 80% ($n_{baseline} = 1,967$) of the Ambit Network's UCLA PTSD-RI data were in the form of the UCLA PTSD-RI-IV, while 20% ($n_{baseline} = 467$) were in the form of the UCLA PTSD-RI-5. In order to use data from both measures, common item equating was used to equate both measures resulting in the "PTSD-RI" dependent variable. In common item equating, Kolen and Brennan (2004) posit that a common item set should be at least 20% of the length of a total test containing 40 or more items. Although the two versions of the UCLA PTSD-RI are shorter than 40 items, 86.4% ($n =$

19 of 22) of the UCLA PTSD-RI-IV items and 70.4% ($n = 19$ of 27) of the UCLA PTSD-RI-5 items overlapped. These proportions exceed Kolen and Brennan's (2004) common item set criterion, suggesting little random equating error.

Common items from the UCLA PTSD-RI-IV and 5 were determined through an item-by-item analysis of both assessment measures (refer to Table 2 for an item-by-item representation of the common items). Specifically, the UCLA PTSD-RI-IV is composed of 22 items across three symptom clusters (i.e., B: Re-Experiencing, C: Avoidance/Numbing, D: Increased Arousal), while the UCLA PTSD-RI-5 is comprised of 27 items across four symptom clusters (i.e., B: Intrusion, C: Avoidance, D: Negative Alterations in Cognition and Mood, and E: Alterations in Arousal and Reactivity). Nineteen items were found to link across UCLA PTSD-RI versions, resulting in a range of 0 to 76 points. Missing PTSD-RI common item data were calculated using mean substitution. At baseline, Cronbach's alphas for the PTSD-RI were .896 for *DSM-IV* and .904 for *DSM-5* common items.

Multilevel modeling. Given the nested nature of the data, multilevel modeling was used to account for the dependency that may result from clustered or non-independent data (Raudenbush & Bryk, 2002). Specifically, multilevel modeling produces unbiased standard error estimates associated with regression coefficients and allows for group characteristics to be included in statistical models of individual outcomes (Raudenbush & Bryk, 2002). The hypothesized model for this study consisted of four levels: timepoints (level 1) were nested within clients (level 2: $n = 2,980$), who were nested within clinicians (level 3: $n = 596$), and clinicians were nested within agencies (level 4: $n = 75$).

Intraclass correlation coefficients. Intraclass correlation coefficients (ICCs) were calculated to determine the proportion of variance in the dependent variables explained by linear mixed model grouping structures. ICCs were calculated using unconditional linear mixed modeling for dependent variables at baseline and follow-up to assess for 2-level (i.e., time-client), 3-level (i.e., time-client-clinician, time-client-agency), and 4-level (i.e., time-client-clinician-agency) clustering of data. Three-level models are referred to as *clinician* (time-client-clinician) and *agency* (time-client-agency) models for ease of reference. Nezlek (2008) estimated that ICCs close to zero are low, while ICCs close to one are high. ICCs for baseline and follow-up assessments are presented in Tables 3 and 4, respectively. When calculating ICCs for the SDQ Self-Report Total at baseline, all 3- and 4-level unconditional models failed to converge. ICCs for the SDQ Parent Report Total scale at baseline were 0.024 for the 3-level clinician model, 0.016 for the 3-level agency model, and 0.015 for the 4-level clinician and 0.012 for the 4-level agency model. Unconditional models did not converge for the PTSD-RI at baseline. At follow-up, ICCs for the SDQ Self-Report Total scale were 0.010 for the 3-level clinician model, 0.016 for the 3-level agency model, and 0.05 for the 4-level clinician and 0.005 for the 4-level agency model. ICCs for the SDQ Parent Report Total scale at baseline were 0.025 for the 3-level agency model, while no other models converged within these parameters. Similar results were found for the PTSD-RI at follow-up (i.e., 0.016 for the 2-level model). Therefore, across baseline and follow-up dependent variables, results suggested that multilevel modeling was not required; however, in order to carry out conservative analyses, multilevel modeling was utilized to account for the clustering of, and missingness, of the data.

Applying multilevel models. First, 4-level (time nested in youth, nested in clinician, nested in agency) mixed models were applied (i.e., random intercept and slope, random intercept only, and random slope only models), but failed to converge. Therefore, 3-level client and agency mixed models were used. Random intercept and slope models failed to converge, thus random intercept only and random slope only models were tested. Although all convergence criteria were satisfied for random slope models, the final Hessian matrix was not positive for all variables except SDQ Self-Report Total. For SDQ Self-Report Total, results from the 3-level random slope model are reported. For SDQ Parent Report Total, SDQ Parent Emotional Symptoms, SDQ Parent Conduct Problems, SDQ Parent Hyperactivity-Inattention, SDQ Parent Prosocial Relations, SDQ Parent Impact, and SDQ Self-Report Impact, the random intercept model converged and results from the analyses are reported. For UCLA PTSD-RI, SDQ Self Emotional Symptoms, SDQ Self-Report Conduct Problems, SDQ Self-Report Hyperactivity-Inattention, SDQ Self-Report Peer Problems, SDQ Parent-Report Peer Problems, and SDQ Self-Report Prosocial Relations variables, neither of the 3-level models converged. Therefore, 2-level mixed models were applied to the outcome variables. Random intercept and slope models were tested, yet did not converge for the remaining variables. Random intercept only and random slope only models were applied, and the random intercept only model converged for UCLA PTSD-RI, SDQ Self-Report Conduct Problems, SDQ Self-Report Peer Problems, SDQ Parent Peer Problems, and SDQ Self-Report Prosocial Relations variables. Results from the random intercept only models are reported. Both random slope only and random intercept only models converged for SDQ Self-Report Emotional Symptoms and SDQ Self-Report Hyperactivity-Inattention

variables. Final models were selected by comparing Akaike information criterion (AIC) and Bayes information criterion (BIC) values (i.e., smaller values indicated better fit), which yielded random intercept only models for SDQ Self-Report Emotional Symptoms and SDQ Self-Report Hyperactivity-Inattention variables.

For all models, means between baseline and follow-up outcomes were the a priori comparison (Time). All models controlled for the number of months between baseline and follow-up assessments (Months in Treatment). When fitting linear mixed models, SPSS uses the Satterthwaite approximation method to compute denominator degrees of freedom. Thus, denominator degrees of freedom used for F statistics are not whole numbers (West, 2009).

Missing data. Missing values were present in the dataset due to the longitudinal nature of the research design. Missing data in all models were managed using the restricted maximum likelihood (REML) procedure used by SPSS 24.0.

Results

Descriptive Characteristics

Complete trauma history data from the GTF were available for 2,980 clients (see Tables 5 and 6). Clients reported an average of 4.0 traumatic events ($SD = 2.9$), ranging from 0 to 15 events. Although 14.5% of clients denied trauma exposure at the time of the baseline assessment ($n = 431$), they were likely referred to trauma treatment given a suspected or confirmed traumatic event. Almost half of clients reported experiencing traumatic loss, bereavement, or separation (45.4%; $n = 1,352$); exposure to an impaired caregiver (45.1%; $n = 1,345$); and witnessing domestic violence (45.1%; $n = 1,343$). Regarding other trauma types, 40.7% of youth reported experiencing emotional abuse or

psychological maltreatment ($n = 1,212$), 35.7% physical abuse ($n = 1063$), 27.2% neglect ($n = 812$), 21.6% sexual assault or rape ($n = 643$), 20.7% sexual abuse ($n = 616$), 19.4% physical assault ($n = 577$), 16.9% school violence ($n = 504$), 15.5% serious accident or injury ($n = 462$), 14.5% extreme interpersonal violence ($n = 432$), 13.3% illness or medical trauma, 12.7% community violence ($n = 377$), and 8.3% natural disaster ($n = 247$). Less than 5% of the sample experienced kidnapping (3.1%; $n = 91$); forced displacement (1.5%; $n = 45$); war, terrorism, or political violence outside the United States (0.9%; $n = 28$); or war, terrorism, or political violence inside the United States (0.5%; $n = 14$).

Client Outcomes

Table 7 presents baseline and follow-up means, standard deviations, and effect sizes for all client outcomes.

PTSD-RI. Clients who participated in TF-CBT reported a significant decrease in posttraumatic stress symptomatology from baseline ($M = 33.17$, $SD = 15.87$) to follow-up ($M = 20.72$, $SD = 14.51$), $F(1, 960.04) = 500.38$, $p = .000$, $d = 0.92$. There was no significant effect of months in treatment on change in PTSD-RI outcome. Estimates of covariance parameters demonstrated significant variability in the mean scores (i.e., intercepts) between baseline and follow-up across youth between baseline and follow-up scores across clients, Wald $Z = 12.85$, $p < .001$. Table 8 presents mixed modeling results.

SDQ Self-Report. For clients participating in TF-CBT, there was an overall significant decrease in total self-reported SDQ score from baseline ($M = 16.63$, $SD = 6.30$) to follow-up ($M = 13.92$, $SD = 6.35$), $F(1, 42.06) = 33.73$, $p < .001$, $d = 0.42$.

Similarly significant reductions from baseline to follow-up were found for self-reported

emotional symptoms, $F(1, 398.32)=44.32, p = .000, d = 0.38$; conduct problems, $F(1, 406.81)=43.44, p < .001, d = 0.33$; hyperactivity-inattention, $F(1, 366.61)=24.28, p < .001, d = 0.27$; peer problems, $F(1, 342.70)=17.79, p < .001, d = 0.23$; and overall impact of difficulties, $F(1, 357.64)=32.33, p < .001, d = 0.27$. There was also a significant increase in self-reported prosocial relations from baseline ($M = 7.61, SD = 2.08$) to follow-up ($M = 7.88, SD = 1.96$), $F(1, 389.55)=5.12, p = .024, d = 0.12$. There was no significant effect of months in treatment on SDQ Self-Report outcomes.

Estimates of covariance parameters demonstrated significant variability in the means scores (i.e., intercepts) between baseline and follow-up scores across clients for the following SDQ Self-Report scales: Emotional Symptoms, Wald $Z = 2.90, p < .001$; Conduct Problems, Wald $Z = 2.25, p < .001$; Hyperactivity-Inattention, Wald $Z = 3.42, p < .001$; Peer Problems, Wald $Z = 2.63, p < .001$; and Prosocial Relations, Wald $Z = 2.19, p < .001$. Notably, estimates of covariance parameters showed significant variability in the mean scores (i.e., intercepts) between clients' baseline and follow-up scores across agencies for the overall impact of difficulties, Wald $Z = 2.91, p < .001$. Table 9 present mixed modeling results for SDQ Self-Report scales.

SDQ Parent Report. For clients participating in TF-CBT, there was an overall significant decrease in total parent-reported SDQ score from baseline ($M = 17.93, SD = 7.39$) to follow-up ($M = 14.66, SD = 7.21$), $F(1, 431.04)=88.59, p < .001, d = 0.49$. Similar significant reductions from baseline to follow-up were discovered for parent-reported emotional symptoms, $F(1, 541.60)=52.88, p < .001, d = 0.46$; conduct problems, $F(1, 409.33)=47.04, p < .001, d = 0.35$; hyperactivity-inattention, $F(1, 419.10)=33.34, p < .001, d = 0.33$; peer problems, $F(1, 435.49)=39.03, p < .001, d = 0.26$; and overall

impact of difficulties, $F(1, 392.07)=80.24, p < .001, d = 0.53$. There was also a significant increase in parent-reported prosocial relations from baseline ($M = 6.93, SD = 2.34$) to follow-up ($M = 7.50, SD = 2.12$), $F(1, 421.37)=14.09, p < .001, d = 0.15$. Months in treatment had no significant impact on changes in parent-reported SDQ scores except for prosocial relations, $F(1, 1,088.84)=-2.51, p = .012$.

Estimates of covariance parameters demonstrated significant variability in the mean scores (i.e., intercepts) between baseline and follow-up scores across clients for peer problems, Wald $Z = 2.86, p < .001$. Further, estimates of covariance parameters demonstrated significant variability in the mean scores (i.e., intercepts) between parental baseline and follow-up scores across agencies for the following SDQ Parent-Report scales: Total Score, Wald $Z = 31.42, p < .001$; Emotional Symptoms, Wald $Z = 3.04, p < .001$; Conduct Problems, Wald $Z = 3.82, p < .001$; Hyperactivity-Inattention, Wald $Z = 4.55, p < .001$; Prosocial Relations, Wald $Z = 3.04, p < .001$; and Impact, Wald $Z = 4.49, p < .001$. Table 10 presents mixed modeling results for SDQ Parent Report scales.

Discussion

The current study investigated the nine-year statewide implementation of TF-CBT by the University of Minnesota's Ambit Network. During this time, 596 Ambit Network-trained clinicians from 75 agencies provided TF-CBT to 2,980 clients. The majority of clients received treatment in community practice settings that offered both outpatient and in-home services (44%), while other settings included inpatient, residential treatment, and correctional agencies. Consistent with research suggesting that most youth in community samples are exposed to at least one traumatic event by late adolescence (i.e., 66%; Copeland, Keeler, Angold, & Costello, 2007), 85.5% of clients in this sample reported

experiencing at least one type of traumatic event, 66.7% reported experiencing three or more types of traumatic events, and 40.8% reported experiencing five or more types of traumatic events. Notably, these trauma exposure rates are higher than rates reported by NCTSN service centers overall where, on average, 80% of clients report experiencing at least one type of traumatic event, 59% report experiencing three or more type of traumatic events, and 31% report experiencing five or more types of traumatic event (Briggs et al., 2013). Thus, results suggest that this sample is more highly traumatized than other community-based and NCTSN service center samples.

Regarding client outcomes, clients participating in TF-CBT reported statistically significant reductions in PTSD symptomatology from baseline to follow-up, as well as self- and parent-reported reductions in emotional symptoms, conduct problems, hyperactivity-inattention, peer problems, and the overall impact of these difficulties from baseline to follow-up. Clients also demonstrated a statistically significant self- and parent-reported improvement in prosocial relations from baseline to follow-up. Effect sizes were largest for PTSD symptomatology (Cohen's $d = 0.92$); parent-reported impact of difficulties (Cohen's $d = 0.52$), total emotional and behavioral difficulties (Cohen's $d = 0.49$), and emotional symptoms (Cohen's $d = 0.46$); as well as self-reported total emotional and behavioral difficulties (Cohen's $d = 0.42$).

These results are consistent with outcomes from a recent TF-CBT implementation study hosted by a Delaware-based community mental health agency contracted by the Delaware Division of Prevention and Behavioral Health Services (Webb et al., 2014). In this study, 12 clinicians provided TF-CBT to 72 youth with study results demonstrating a large effect size for PTSD symptoms (Cohen's $d = 0.92$), medium to large effect size for

internalizing symptoms (Cohen's $d = 0.67$), and small to medium effect size for externalizing symptoms (Cohen's $d = 0.39$) (Webb et al., 2014).

Therefore, results from this study are strikingly similar to effect sizes reported by Webb and colleagues (2014) in a smaller TF-CBT implementation study. The magnitude of these outcomes is significant considering that clients in this sample typically presented with variable target problems (e.g., behavioral dysregulation, internalizing distress), exposure to an average of four traumatic events, and heterogeneous trauma histories. Moreover, learning collaborative clinicians provided TF-CBT from diverse professional backgrounds (i.e., psychology, social work, counseling) and implemented TF-CBT among heterogeneous community practice settings (i.e., outpatient, inpatient, home-based, and correctional facilities). In short, results suggest that TF-CBT can be transported to community practice settings and that clinicians can be trained to effectively and potently treat trauma-exposed youth.

Additionally, findings suggest that clinicians in community practice settings are increasingly able to adopt components of trauma-focused EBTs, including the utilization of empirically supported assessment methods (Garland, Kruse, & Aarons, 2003). Within mental health service delivery, for example, there has been increased pressure to collect standardized outcome data on clients served, and such data are now required by many public and private agencies (Garland et al., 2003). In fact, regular outcome assessment within community based practice settings can be regarded as one of the most direct links between research and practice as outcome data can be used to immediately improve service provision (Knaup, Koesters, Schoefer, Becker, & Puschner, 2009). Specific to TF-CBT, robust outcome monitoring not only allows the clinician to better understand

the child or adolescent's trauma history and posttraumatic stress symptomatology, but it also provides the clinician with the opportunity to begin gradual exposure to trauma-related details (Cohen et al., 2017).

This study has several significant limitations common to applied, community-based research. First, there was no comparison or control group, which is common among large-scale implementation and dissemination efforts (see Lang et al., 2015; Sigel et al., 2013). Lack of a control group precludes the ability to make causal inference from the data, including the determination of whether improvements in posttraumatic stress, behavioral problems, and emotional symptoms may be attributable to effects of other treatments received or the passage of time (Webb et al., 2014). Additionally, given the long-term nature of this study, there was a substantial amount of missing outcomes data. It has long been recognized that incomplete, or missing, data are not unusual in longitudinal research designs as many participants are not available at all time points (Ibrahim & Molenberghs, 2009).

For example, in a randomized implementation study of Alternatives for Families: A Cognitive-Behavioral Therapy (AF-CBT), the investigators reported that approximately 50% of practitioners had some missing data for at least one time point (Kolko et al., 2012). Within the Ambit Network's learning collaborative, clinicians were required to administer evidence-based assessments at baseline and every three months until clients completed, or terminated, TF-CBT. The amount of missing data at follow-up may point to the fact that once clinicians graduated from the learning collaborative, they were less likely to administer and/or submit completed assessments. Additional TF-CBT implementation research is needed to examine factors related to the sustainability of not

only the TF-CBT treatment model itself, but other practices like systematic outcome assessment.

Despite the challenges of implementation research, the present study indicates that the large-scale implementation of TF-CBT is a promising approach in increasing the availability of trauma-focused EBTs with heterogenous client populations in diverse community practice settings. These findings have important implications as state- and federal-level policymakers dedicate more funding to the implementation of trauma-focused EBTs to improve the lives of youth, families, and communities affected by trauma.

Table 1

Client Demographic Characteristics (N = 2,980)

	<u>μ</u>	<u>SD</u>
Age	13.0	3.1
	<u>n</u>	<u>%</u>
Gender		
Male	1,296	43.5
Female	1,684	56.5
Race		
Caucasian	1,804	60.5
Black/African American	355	11.9
American Indian	241	8.1
Asian	40	1.3
Native Hawaiian/Pacific Islander	12	0.4
Multiracial	288	9.7
Unknown	240	8.1
Ethnicity		
Hispanic	282	9.5
Primary Residence		
Home with Parents	1,664	55.8
Residential Treatment Center	480	16.1
Kinship Care	312	10.5
Foster Care	284	9.5
Traditional Foster Care	246	8.3
Treatment Foster Care	38	1.3
Correctional Facility	23	0.8
Other	216	7.2

Table 2

UCLA PTSD-RI for DSM-IV and DSM-5 Item Comparisons: Preparation for Common Item Equating

<u>UCLA PTSD-RI for DSM-IV</u>	<u>UCLA PTSD-RI for DSM-5</u>
1. (D4) I watch out for danger or things that I am afraid of.	1. (E3) I am on the lookout for danger or things that I am afraid of.
2. (B4) When something reminds me of what happened, I get very upset, afraid, or sad.	11. (B4) When something reminds me of what happened I get very upset, afraid, or sad.
3. (B1) I have upsetting thoughts, pictures, or sounds of what happened come into my mind when I do not want them to.	18. (B1) I have upsetting thoughts, pictures or sounds of what happened come into my mind when I don't want them to.
4. (D2) I feel grouchy, angry, or mad.	
5. (B2) I have dreams about what happened or other bad dreams.	10. (B2) I have bad dreams about what happened, or other bad dreams.
6. (B3) I feel like I am back at the time the bad thing happened, living through it again.	5. (B3) I feel like I am back at the time when the bad things happened, like it's happening all over again.
7. (C4) I feel like staying by myself and not being with my friends.	7. (D5) I don't feel like doing things with my family or friends or other things that I liked to do.
8. (C5) I feel alone inside and not close to other people.	17. (D6) I feel alone even when I am around other people
9. (C1) I try not to talk about, think about, or have feelings about what happened.	13. (C1) I try not to think about or have feelings about what happened.
10. (C6) I have trouble feeling happiness or love.	12. (D7) I have trouble feeling happiness or love.
11. (C6) I have trouble feeling sadness or anger.	
12. (D5) I feel jumpy or startle easily, like when I hear a loud noise or when something surprises me.	24. (E4) I feel jumpy or startle easily, like when I hear a loud noise or when something surprises me.
13. (D1) I have trouble going to sleep or I wake up often during the night.	21. (E6) I have trouble going to sleep, wake up often, or have trouble getting back to sleep.
14. (AF) I think that some part of what happened is my fault.	19. (D3) I feel that part of what happened was my fault.
15. (C3) I have trouble remembering important parts of what happened.	23. (D1) I have trouble remembering important parts of what happened.
16. (D3) I have trouble concentrating or paying attention.	8. (E5) I have trouble concentrating or paying attention.
17. (C2) I try to stay away from people, places, or things that make me remember what happened.	3. (C2) I try to stay away from people, places, or things that remind me about what happened.

18. (B5) When something reminds me of what happened, I have strong feelings in my body, like my heart beats fast, my head aches, or my stomach aches.	14. (B5) When something reminds me of what happened, I have strong feelings in my body like my heart beats fast, my head aches, or my stomach aches.
19. (C7) I think that I will not live a long life.	
20. (D2) I have arguments or physical fights.	4. (E1) I get upset easily or get into arguments or physical fights.
21. (C7) I feel pessimistic or negative about my future.	9. (D2) I have thoughts like, "The world is really dangerous."
22. (AF) I am afraid that the bad thing will happen again.	25. (D4) I feel afraid or scared.
	2. (D2) I have thoughts like "I am bad."
	15. (D3) I am mad with someone for making the bad thing happen, not doing more to stop it, or to help after.
	16. (D2) I have thoughts like, "I will never be able to trust other people."
	20. (E2) I hurt myself on purpose.
	22. (D4) I feel ashamed or embarrassed over what happened.
	26. (E2) I do risky or unsafe things that could really hurt me or someone else.
	27. (D4) I want to get back at someone for what happened.

Note. Letters and numbers in parentheses correspond to the symptom cluster and number, respectively. For the UCLA PTSD-RI for *DSM-IV* symptoms clusters include: B = Re-Experiencing, C = Avoidance/Numbing, and D = Arousal. For the UCLA PTSD-RI for *DSM-5* symptom clusters include: B = Intrusion Symptoms, C = Avoidance of Stimuli, D = Negative Alterations in Cognitions and Mood, and E = Marked Alterations in Arousal and Reactivity.

Table 3

Intraclass Correlation Coefficients for Client Outcome Variables at Baseline

	3-Level (Clinician, <u>Agency</u>)	2-Level (<u>Clinician</u>)	2-Level (<u>Agency</u>)
PTSD-RI	--	--	--
SDQ Self Report			
Total	--	--	--
Emotional Symptoms	--	--	--
Conduct Problems	Clinician = .009 Agency = .004	.003	.006
Hyperactivity-Inattention	--	.009	--
Peer Relationship Problems	--	--	--
Prosocial Behavior	--	.090	.002
Impact	--	.010	--
SDQ Parent Report			
Total	Clinician = .015 Agency = .012	.024	.016
Emotional Symptoms	--	--	.003
Conduct Problems	--	.011	.011
Hyperactivity-Inattention	--	.030	.02
Peer Relationship Problems	--	--	--
Prosocial Behavior	--	.007	.009
Impact	Clinician = .05 Agency = .03	.060	.040

Note. -- = no convergence

Table 4

Intraclass Correlation Coefficients for Client Outcome Variables at Follow-Up

	3-Level (Clinician, <u>Agency</u>)	2-Level (<u>Clinician</u>)	2-Level (<u>Agency</u>)
PTSD-RI	--	.016	--
SDQ Self Report			
Total	Agency = .005 Clinician = .050	.010	.016
Emotional	--	--	--
Conduct Problems	--	--	--
Hyperarousal-Inattention	--	.003	--
Peer	--	--	--
Prosocial	Agency = .023 Clinician = .086	.087	.045
Impact	Agency = .054 Clinician = .095	.121	.084
SDQ Parent Report			
Total	--	--	.025
Emotional	Agency = .006 Clinician = .018	.008	.006
Conduct Problems	--	.020	.042
Hyperarousal-Inattention	--	--	--
Peer	Agency = .007 Clinician = .015	.027	.010
Prosocial	--	.058	--
Impact	--	--	.034

Note. -- = no convergence

Table 5

*Mean and Total Number of Traumas Reported
by Clients (N = 2,980)*

	<u>μ</u>	<u>SD</u>
Mean Number	4.0	2.9
	<u>n</u>	<u>%</u>
Total Number		
0	431	14.5
1	224	7.5
2	336	11.3
3	346	11.6
4	427	14.3
5	340	11.4
6	310	10.4
7	211	7.1
8	160	5.4
9	86	2.9
10	55	1.8
11	21	0.7
12	23	0.8
13	5	0.2
14	4	0.1
15	1	0.0

Table 6

Proportion of Sample Endorsing Specific Trauma Types (N = 2, 980)

	<i>n</i>	<i>%</i>
Traumatic loss/bereavement/separation	1,352	45.4
Impaired caregiver	1,345	45.1
Domestic violence	1,343	45.1
Emotional abuse/psychological maltreatment	1,212	40.7
Physical abuse/maltreatment	1,063	35.7
Neglect	812	27.2
Sexual assault/rape	643	21.6
Sexual abuse/maltreatment	616	20.7
Physical assault	577	19.4
School violence	504	16.9
Serious accident/injury	462	15.5
Extreme interpersonal violence	432	14.5
Illness/medical trauma	396	13.3
Community violence	377	12.7
Natural disaster	247	8.3
Kidnapping	91	3.1
Forced Displacement	45	1.5
War/Terrorism/Political Violence Outside the US	28	0.9
War/Terrorism/Political Violence Inside the US	14	0.5
Other	367	12.3

Note. Trauma types are not mutually exclusive

Table 7

Client Outcomes at Baseline and Follow-Up

	<i>n</i> (B/F)	Baseline μ (SD)	Follow-Up μ (SD)	Effect Size (<i>d</i>)
PTSD-RI	2,495/621	33.17 (15.87)	20.72 (14.51)	0.92
SDQ Self-Report				
Total	769/274	16.63 (6.30)	13.92 (6.35)	0.42
Emotional Symptoms	770/276	4.70 (2.59)	3.81 (2.61)	0.38
Conduct Problems	770/275	3.28 (2.07)	2.46 (1.96)	0.33
Hyperactivity-Inattention	770/276	5.51 (2.39)	4.88 (2.41)	0.27
Peer Problems	770/276	3.23 (2.18)	2.72 (1.84)	0.23
Prosocial Relations	770/276	7.61 (2.08)	7.88 (1.96)	0.12
Impact	713/259	2.52 (2.58)	1.53 (2.14)	0.27
SDQ Parent Report				
Total	999/309	17.93 (7.39)	14.66 (7.21)	0.49
Emotional Symptoms	1,001/310	4.75 (2.65)	3.78 (2.56)	0.46
Conduct Problems	1,002/311	3.91 (2.60)	3.00 (2.42)	0.35
Hyperactivity-Inattention	1,002/311	5.91 (2.84)	5.24 (2.65)	0.33
Peer Problems	1,001/311	3.43 (2.23)	2.65 (2.08)	0.26
Prosocial Relations	1,001/311	6.93 (2.34)	7.50 (2.12)	0.15
Impact	939/287	3.76 (2.93)	2.34 (2.63)	0.52

Note. B: Baseline; F: Last Follow Up or Closing; UCLA PTSD-RI: UCLA

Posttraumatic Stress Disorder Reaction Index; SDQ: Strengths and Difficulties

Questionnaire; *d*: Cohen's *d*.

Table 8

PTSD-RI: Estimated Fixed Effects and Covariance Parameters

	β	$SE(\beta)$	t	p
Time	-13.01	0.58	-22.37	.000
Months in Treatment	0.01	0.05	0.29	.771
Estimates of Covariance				
	<u>Estimate</u>	<u>SE</u>	<u>Wald Z</u>	p
Client	118.32	9.21	12.85	.000

Table 9

SDQ Self-Report: Estimated Fixed Effects and Covariance Parameters

Fixed Effects				
	β	SE(β)	t	p
Total Score				
Time	-2.71	0.47	-5.81	.000
Months in Treatment	-0.00	0.03	-0.08	.933
Emotional Symptoms				
Time	-0.95	0.14	-6.66	.000
Months in Treatment	-0.01	0.01	-0.97	.334
Conduct Problems				
Time	-0.76	0.12	-6.59	.000
Months in Treatment	0.01	0.01	0.98	.326
Hyperactivity-Inattention				
Time	-0.64	0.13	-4.93	.000
Months in Treatment	-0.00	0.01	-0.68	.498
Peer Problems				
Time	-0.49	0.12	-4.22	.000
Months in Treatment	0.00	0.01	0.11	.92
Prosocial Relations				
Time	0.27	0.12	2.26	0.02
Months in Treatment	-0.01	0.01	-0.76	.447
Impact				
Time	-0.87	0.15	-5.69	.000
Months in Treatment	-0.02	0.01	-1.42	0.16
Estimates of Covariance				
	Estimate	SE	Wald Z	p
Total Score				
Agency	0.64	1.57	0.41	.684
Client*Agency	0.04	4.18	0.10	.992
Emotional Symptoms				
Client	2.90	0.28	10.41	.000
Conduct Problems				
Client	2.25	0.24	9.32	.000
Hyperactivity-Inattention				
Client	3.42	0.33	10.34	.000
Peer Problems				
Client	2.63	0.27	9.73	.000
Prosocial Relations				
Client	2.19	0.26	8.32	.000
Impact				
Agency	0.04	0.95	0.41	.682
Client*Agency	2.91	0.45	6.44	.000

Table 10

SDQ Parent-Report: Estimated Fixed Effects and Covariance Parameters

Fixed Effects				
	β	$SE(\beta)$	t	p
Total Score				
Time	-3.47	0.7	-9.41	.000
Months in Treatment	0.06	0.04	1.56	.119
Emotional Symptoms				
Time	-1.07	0.15	-7.27	.000
Months in Treatment	0.00	0.13	0.34	.731
Conduct Problems				
Time	-0.89	0.13	-6.86	.000
Months in Treatment	0.02	0.01	1.94	.053
Hyperactivity-Inattention				
Time	-0.80	0.14	-5.77	.000
Months in Treatment	0.02	0.01	1.73	.083
Peer Problems				
Time	-0.69	0.11	-6.25	.000
Months in Treatment	0.01	0.01	0.97	.334
Prosocial Relations				
Time	0.44	0.12	3.75	.000
Months in Treatment	-0.03	0.01	-2.51	.012
Impact				
Time	-1.39	0.15	-8.96	.000
Months in Treatment	-0.01	0.01	-0.11	.911
Estimates of Covariance				
	<u>Estimate</u>	<u>SE</u>	<u>Wald Z</u>	p
Total Score				
Agency	1.28	0.76	1.69	.091
Client*Agency	31.42	2.77	11.36	.000
Emotional Symptoms				
Agency	0.02	0.06	0.36	.722
Client*Agency	3.04	0.40	7.68	.000
Conduct Problems				
Agency	0.10	0.07	1.48	.139
Client*Agency	3.82	0.35	10.80	.000
Hyperactivity-Inattention				
Agency	0.24	0.13	1.84	.065
Client*Agency	4.55	0.41	11.23	.000
Peer Problems				
Client	2.86	0.25	11.50	.000
Prosocial Relations				
Agency	0.04	0.05	0.76	.447
Client*Agency	3.04	0.28	10.83	.000
Impact				

Agency	0.31	0.17	1.86	.063
Client*Agency	4.49	0.47	9.45	.000

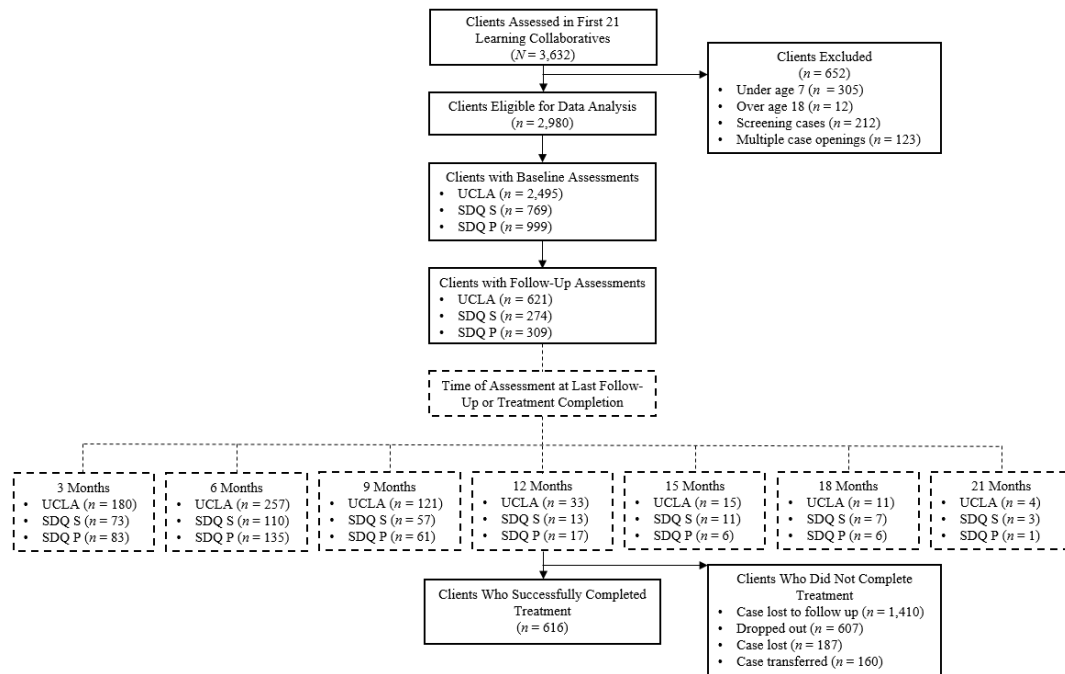


Figure 1. Flow diagram of eligible clients and associated baseline and follow-up assessments.

Note. UCLA: UCLA Posttraumatic Stress Disorder—Reaction Index for *DSM-IV* and *DSM-5*; SDQ S: Strengths and Difficulties Questionnaire, Self-Report Version; SDQ P: Strengths and Difficulties Questionnaire, Parent Report Version; screening cases: as part of an early learning collaborative, Ambit Network trained clinicians to screen for posttraumatic stress symptomatology using the UCLA, although clinicians were not required to conduct follow-up assessments; multiple case openings: for clients who began Trauma-Focused Cognitive-Behavioral Therapy more than once, a decision tree was utilized to select the appropriate course of therapy to be used in data analysis.

**The Large-Scale Implementation of Trauma-Focused Cognitive-Behavioral
Therapy (TF-CBT) in Community Practice Settings:
An Assessment of Implementation Fidelity**

Within the last decade, considerable advances have been made in the dissemination and implementation of evidence-based treatments (EBTs) for trauma-exposed youth in community practice settings (Kolko et al., 2012). Indeed, national (Pynoos et al., 2008) and state-level (Lang et al., 2017; Sigel et al., 2013) initiatives have established or enhanced large-scale efforts to train community-based clinicians in trauma-focused EBTs to reduce the detrimental public health impact of child abuse and neglect. In spite of this growth, however, the penetration rate for EBTs in community practice settings remains limited and is estimated to be only one to three percent (Bruns et al., 2015). As such, the continued investigation of large-scale dissemination and implementation efforts for trauma-focused EBTs is critical to ensure their effectiveness for heterogeneous client populations and transportability to a variety of community practice settings where trauma-exposed youth are primarily served.

One of the most broadly disseminated and implemented EBTs for trauma-exposed youth in the United States is Trauma-Focused Cognitive-Behavioral Therapy (TF-CBT; Cary & McMillen, 2014; Cohen et al., 2017). TF-CBT is a components-based treatment model for youth ages 3 to 18 and their caregivers (Cohen et al., 2006), and has been tailored for use with youth exposed to a variety of traumatic experiences (Allen & Johnson, 2012; Cohen et al., 2006; Cohen et al., 2012), from heterogeneous cultural backgrounds (Bigfoot & Schmidt, 2010; Murray, Familiar et al., 2013), and served in both individual and group treatment settings (Deblinger, Pollio, & Dorsey, 2015;

McMullen, O’Callaghan, Shannon, Black, & Eakin, 2013; O’Callaghan, McMullen, Shannon, Rafferty, & Black, 2013). Although a large body of research has documented TF-CBT’s efficacy and effectiveness, TF-CBT’s developers and purveyors are only beginning to examine the process of transporting, or implementing, TF-CBT on a larger scale and in routine mental health care settings.

To facilitate the implementation of TF-CBT, the National Child Traumatic Stress Network (NCTSN) adapted the Institute for Healthcare Improvement’s Breakthrough Series Collaborative (BSC) model to structure and streamline the implementation of TF-CBT across the United States (Ebert et al., 2012). The BSC is a package of implementation strategies designed to cultivate child-service system buy-in and to enhance the long-term sustainability of the TF-CBT model (Sigel et al., 2013, p. 324). Group-level training is emphasized and consists of separate training tracks for clinicians, supervisors, and senior leaders (i.e., administrators and directors) to enhance organizational change (Ebert et al., 2012; Ebert, Amaya-Jackson, Markiewicz, & Burroughs, 2008). Over the course of one year, staff within each training track attend several in-person trainings and participate in ongoing phone-based or in-person consultation (Ebert et al., 2012; Ebert et al., 2008). Consultation generally targets different aspects of the TF-CBT treatment model, as well as effective implementation strategies and broader dissemination issues (Sigel et al., 2013).

Since 2005, multiple states have utilized the BSC to implement TF-CBT across community practice settings (e.g., Sigel et al., 2013; Lang et al., 2017). In Connecticut, for example, the Connecticut Department of Children and Families, chose to deliver and sustain TF-CBT throughout outpatient children’s behavioral health clinics due to

increased awareness of the needs of trauma-exposed children in the child welfare system (Greeson et al., 2001; Lang et al., 2015). Using the BSC over the course of nine years, the Connecticut Department of Children and Families trained a total of 612 clinicians from 87 agencies in TF-CBT, serving approximately 6,200 youth (Lang et al., 2017). Connecticut's TF-CBT implementation initiative also examined the quality of TF-CBT delivery across the state by assessing engagement, treatment completion of, and fidelity to the TF-CBT model. Engagement, which was operationalized as the completion of at least four TF-CBT sessions, was estimated to be 81.7% among clients, while 33.9% of clients successfully completed TF-CBT. Implementation fidelity was defined as the percent of required TF-CBT components completed, and was estimated to be 89.0% among client-clinician pairs successfully completing the TF-CBT model. Although Connecticut's statewide implementation of the BSC represents an impressive addition to the TF-CBT dissemination and implementation literature, it proffered simplistic examinations of implementation outcomes, most notably, implementation fidelity.

Implementation Fidelity

Implementation fidelity is defined as the degree to which a clinician delivers a treatment adherent to its core principles (Carroll et al., 2007; Proctor et al., 2011) or, comparably, the degree to which treatment models are delivered as intended by the treatment developers (Proctor et al., 2011). Within the dissemination and implementation science literature, fidelity has been measured more often than any other implementation outcome (e.g., adoption, feasibility, sustainability) given that insufficient fidelity to EBTs often directly affects the magnitude of intervention outcomes. Therefore, when large-scale implementation initiatives fail, as they often do, the careful examination of

implementation fidelity allows for the ascertainment of whether the failure occurred because the EBT was ineffective in the new setting or if an appropriate EBT was deployed incorrectly (Proctor et al., 201).

Notably, there exists a lack of consensus among implementation researchers as to what specifically constitutes implementation fidelity (Dane & Schneider, 1998). For instance, some experts proffer that implementation fidelity is composed of three components (i.e., treatment adherence, treatment differentiation, and treatment competence; Schoenwald et al., 2011), while others suggest up to eight (i.e., adherence, dosage, quality, participant responsiveness, program differentiation, monitoring of control/comparison conditions, program reach, and adaptation; Durlak & DuPre, 2008). As described in their conceptual framework for implementation fidelity, Carroll and colleagues (2007) identified four primary components of implementation fidelity, including the content, frequency, duration, and coverage (i.e., dose) of an intervention.

Specifically, an intervention's *content* is generally considered to be its core components, or the skills, knowledge, or elements of treatment that must be delivered to participants. The *frequency* of implementation is characterized as the intervention's rate of occurrence (e.g., daily, weekly, biweekly), while *duration* refers to length of intervention delivery (e.g., total number of sessions). Finally, *coverage* measures the extent to which "people who should be participating in or receiving the benefits of an intervention actually do so" (p. 41). Given the demonstrated overlap between measures of adherence and quality of delivery, as well as the widespread recognition of Carroll et al.'s (2007) framework for implementation fidelity, this study operationalizes implementation fidelity as consisting of the content, frequency, duration, and coverage of an intervention.

Researchers across disciplines have begun to investigate the impact of implementation fidelity on client outcomes. For example, a 2008 review of five meta-analyses investigated the impact of various prevention and intervention programs (e.g., anti-bullying programs, drug use prevention studies, school-based prevention programs targeting aggressive behaviors) on client outcomes (Durlak & DuPre, 2008). Results revealed that interventions with better, as opposed to poorer, implementation demonstrated effect sizes that were two to three times larger. Under ideal circumstances, these effect sizes increased to twelve times larger with better implementation fidelity (Derzon, Sale, Springer, & Brounstein, 2005). Of course, such review and meta-analytic findings are limited by the data presented in the original investigations, particularly with respect to the precise level of implementation achieved (Durlak & DuPre, 2008). To date, no large-scale TF-CBT implementation study has directly investigated the relationship between implementation fidelity and client outcomes in community practice settings.

Factors Associated with Implementation Fidelity

Due to the complexity of implementing an EBT in a fidelitous manner, a variety of factors are estimated to influence implementation fidelity and these factors occur across multiple levels of the practice context, including agency (Aarons et al., 2009; Fixsen et al., 2005), clinician (Durlak & DuPre, 2008; Fixsen et al., 2005; Kolko et al., 2009; Ruzek & Rosen, 2009), and client (Fixsen et al., 2005) levels. Within the TF-CBT implementation literature, no research has investigated the client and agency characteristics that enhance or diminish TF-CBT implementation fidelity and, surprisingly, only a handful of researchers have examined the clinician-level factors that

are associated with TF-CBT implementation fidelity (see Allen & Johnson, 2008; Cohen et al., 2016; Fritz et al., 2013; Lang et al., 2015; Woody, Anderson, & D'Souza, 2015);

While agency-level factors have not been assessed within TF-CBT implementation fidelity research, Aarons and colleagues (Aarons, Hurlburt, & Horwitz, 2011) suggest that agencies with greater absorptive capacity—or preexisting knowledge, ability to use knowledge, specialization, and means to encourage knowledge sharing—are more likely to implement EBTs with fidelity. Consequently, public sector agencies, such as child welfare or juvenile justice, have been shown to demonstrate less absorptive capacity due to workforces with varied levels of education, heavy caseloads, and the delivery of varied services (Yoo, Brooks, & Patti, 2007).

Moreover, significant variability exists in the extent to which clinician-level characteristics affect implementation fidelity. Clinicians who are responsible for delivering trauma-focused EBTs, for example, may differ by level of experience (Rohrbach, Graham, & Hansen, 1993), as well as motivation and self-efficacy (Rohrbach et al., 1993; Mihalic, Fagan, & Argamaso, 2008). Within the TF-CBT implementation literature, Allen and Johnson (2012) discovered that clinicians' professional discipline (i.e., counseling, social work, clinical psychology) was not related to the likelihood of regularly implementing all TF-CBT components. This result was replicated by Cohen and colleagues (2016), in that clinicians' professional discipline did not significantly impact TF-CBT implementation outcomes. Similarly, no discipline differences (e.g., degree, type of license) were found to exist between participants and nonparticipants in TF-CBT consultation calls (i.e., as studied in AR BEST's model; Fritz et al., 2013). Regarding licensure status, licensed therapists have been found to be significantly more likely than

non-licensed therapists to assess for PTSD and depression symptoms, engage and retain youth in TF-CBT treatment, and to complete TF-CBT with fidelity, with a trend for licensure status to impact completion of the model (Cohen et al., 2016).

Research is only beginning to systematically evaluate client demographic and trauma characteristics that may affect implementation fidelity to traditional EBTs (Chaudoir, Dugan, & Barr, 2013), let alone trauma-focused EBTs. Client-level characteristics of interest include the number of traumas experienced, onset of first trauma, type of trauma experienced, and other demographic characteristics such as race, ethnicity, gender, or place of residence. Results from a study examining the implementation of Cognitive Therapy for PTSD suggest that adult clients treated for multiple traumas were less likely to attend sessions regularly compared with clients who were treated for one or two traumas (Ehlers et al., 2013). Ehlers et al. (2013) also discovered that the treatment for clients with multiple traumas was less trauma-focused than the treatment of other clients, although the authors did not find statistical differences in dropout rate or total number of treatment sessions completed.

Current Study

The purpose of this study is to examine the relationship between implementation fidelity and client outcomes using data from a statewide TF-CBT implementation initiative that trained 596 clinicians across 75 agencies, and served 2,980 youth. In addition to the association between fidelity and outcomes, this investigation also assesses agency-, clinician-, and client-factors related to TF-CBT implementation fidelity. Based on Carroll et al.'s research on implementation science, TF-CBT implementation fidelity is conceptualized as a multidimensional construct comprised of four measures: 1) content

(i.e., delivery of required TF-CBT components), 2) frequency (i.e., rate of TF-CBT session occurrence, 3) duration (i.e., number of total TF-CBT sessions), and 4) coverage (i.e., completion of TF-CBT as planned).

Presenting data from a large-scale implementation initiative, this study seeks to answer two questions. First, is clinician implementation fidelity to the TF-CBT treatment model associated with posttraumatic stress, emotional problem, and behavioral problem (i.e., conduct problems, hyperactivity-inattention, peer problems) change scores from baseline to follow-up in a community sample of clients? It is predicted that clinician implementation fidelity will be positively associated with posttraumatic stress, emotional problem, and behavioral problem change scores from baseline to follow-up.

Second, what client-, clinician-, and organization-level factors are associated with implementation fidelity to the TF-CBT treatment model? In line with previous research on agency-level factors, it is hypothesized that type of agency and number of agency sites (i.e., serving as proxies for absorptive capacity) will be associated with TF-CBT implementation fidelity. Regarding clinician-level factors, clinician supervisor status (i.e., serving as a proxy for level of experience) is posited to be associated with TF-CBT implementation fidelity, while for client-level factors, it is estimated that the number of traumatic events experienced will be associated with implementing TF-CBT in a fidelitous manner.

Method

Implementation Process

Background. From 2007 to 2016, the University of Minnesota’s Ambit Network implemented over 21 learning collaboratives—akin to NCTSN’s BSCs—to train mental health clinicians in TF-CBT. In order to maintain confidentiality, assessment measures

were deidentified by clinicians before they were processed by the Ambit Network. Fidelity measures did not include any type of private or confidential client data.

Selection process. Seventeen (81.0%) learning collaboratives were co-sponsored by Minnesota DHS, while four (19.0%) were independently sponsored by the Ambit Network.

Minnesota DHS agency selection. Agencies were identified and recruited by Minnesota DHS through requests for proposals (RFP). The purpose of the RFP was to develop “clinical capacity within mental health provider agencies to effectively utilize existing and emerging research to inform all treatment, and improve outcomes for children and their families” (Minnesota DHS & Children’s Mental Health Division, 2008, p. 3). Agencies were required to apply as teams, with one clinician serving as a clinical supervisor, and two to three clinicians serving as direct service providers. Agencies selected to participate in the learning collaborative were contracted by Minnesota DHS for 18 months. These contracts provided \$15,000 per clinician to cover the costs associated with participating in the learning collaborative, including time the clinician spent in online training, in-person training, on bi-monthly phone consultation, conducting assessments not covered by insurance, and fidelity monitoring. In addition, contracts also reimbursed tuition, lodging, and per diem costs associated with traveling to training locations.

Ambit Network agency selection. Agencies were also recruited by the Ambit Network through an in-depth application process. This process was comprised of a ten-page application, which assessed each agency’s previous experience with manualized interventions, fidelity monitoring, administrative support to implement new practices and

provide time for supervision, among other areas. Preference was given to sites that wished to train more than one therapist under the purview of a primary supervisor. Clinicians were charged \$1,000 to \$1,200 to participate, depending on demand.

Learning collaborative model. Before in-person training began, all clinicians were required to complete *TF-CBTWeb*, a ten-hour online learning course for TF-CBT (<https://tfcbt.musc.edu>; Medical University of South Carolina, 2005). All trainings were conducted by an approved TF-CBT national trainer (i.e., graduate of the TF-CBT Train-the-Trainer Program) and included a variety of active-learning techniques (e.g., assessment role plays, case vignettes, small group discussions, interactive games, and other strategies; Lang et al., 2015; Lyon, Stirman, Kerns, & Bruns, 2011). Approximately one-third of the training focused on enhancing clinical skills required to assess child traumatic stress, one-third focused on learning the TF-CBT treatment model, and one-third focused on methods of implementation and data collection through trauma-informed assessments. Clinicians were also required to participate in biweekly consultation calls with a TF-CBT consultant and their cohort mates, and presented at least one case for consultation with the TF-CBT consultant. Designated supervisors were also required to participate in monthly consultation calls.

Intervention

Trauma-Focused Cognitive-Behavioral Therapy (TF-CBT) is a components-based treatment model for youth ages 3 to 18 and their caregivers affected by trauma (Cohen et al., 2006). Efficacy and effectiveness trials have concluded that children participating in TF-CBT demonstrate improvements in posttraumatic stress disorder (PTSD), depression, and anxiety, externalizing behavior problems, relationship and attachment problems, and

academic difficulties (Cohen et al., 2006), while parents demonstrate reduced parental distress and improvements in parenting skills (for reviews, see Cary & McMillen, 2012; Cohen et al., 2017). The TF-CBT model is comprised of eight essential elements: Psychoeducation and Parenting Skills (P), Relaxation (R), Affect Expression and Regulation (A), Cognitive Coping (C), Trauma Narrative Development and Processing (T), In Vivo Gradual Exposure (I), Conjoint Parent/Child Sessions (C), and Enhancing Safety/Future Development (E). Summarized by the acronym PRACTICE, TF-CBT's components allow for a flexible, manualized treatment approach. The treatment model consists of 12 to 25 weekly sessions, which are generally 60 to 90 minutes in length and are conducted on a weekly basis at a pace determined by the youth's affective and cognitive readiness (Cohen et al., 2006).

Participants

Agencies. Seventy-five agencies were selected to participate in the learning collaboratives. The majority of agencies were housed at one location ($n = 34$; 45.3%) and offered both outpatient and in-home mental health services ($n = 33$; 44.0). The remaining agencies provided outpatient only ($n = 13$, 17.3%); outpatient, in-home, and inpatient ($n = 11$, 14.7%); outpatient and inpatient ($n = 7$, 9.3%); inpatient only ($n = 5$, 6.7%); comprehensive (i.e., outpatient, in-home, inpatient, and correction/residential treatment; $n = 3$, 4.0%); correction/residential treatment only ($n = 2$, 2.7%); and in-home only ($n = 1$, 1.3%) services. Agency characteristics are presented in Table 11.

Clinicians. A total of 596 clinicians participated in the learning collaborative models, of which 105 were supervisors (17.6%). Clinicians predominantly identified as female ($n = 490$; 82.2%). Regarding training, the majority of clinicians had a Masters in

Social Work ($n = 206$; 34.6%), Masters of Arts ($n = 202$; 33.9%), or Masters of Science ($n = 98$; 16.4%) and were licensed in social work ($n = 211$; 35.4%), marriage and family therapy ($n = 143$; 24.0%), and professional counseling ($n = 124$; 20.8%). Sixty-three clinicians were trained at the doctoral level, either in psychology ($n = 33$; 5.5%) or philosophy ($n = 30$; 5.0%) and licensed as psychologists ($n = 48$; 8.1%). Clinician characteristics are presented in Table 12.

Clients. Although clients as young as 3 years of age participated in TF-CBT, the study sample was limited to clients between the ages of 7 and 18 given age cutoffs for the UCLA PTSD-Reaction Index for *DSM-IV* (American Psychiatric Association, 2000) and *DSM-5* (American Psychiatric Association, 2013). Specifically, 305 clients were excluded from analyses as they were younger than age 7, and 12 clients were excluded as they were older than age 18. Analyses also excluded 212 clients who had participated in an early TF-CBT learning collaborative (i.e., clients were only administered the UCLA PTSD-Reaction Index for *DSM-IV* without the collection of associated demographic and fidelity data), as well as 123 clients with multiple case openings. For clients with multiple case openings, the author used a decision tree to select the most recent and/or complete case opening. Using the aforementioned criteria, 2,980 clients between the ages of 7 and 18 completed an assessment to determine the appropriateness of TF-CBT. In accordance with the Strengthening the Reporting of Observational Studies in Epidemiology Statement (von Elm et al., 2007), refer to Figure 1 for a flow diagram of eligible clients and associated baseline and follow-up assessments.

Clients presented for treatment to the agencies for a range of concerns, and were identified by the clinicians as appropriate for TF-CBT during screening and assessment. Client demographics are presented in Table 1. Clients were an average of 12.96 years of age ($SD = 3.14$, $Range = 7.0 - 21.3$), and primarily identified as female (56.5%; $n = 1,684$). The racial and ethnic makeup of the clients was diverse. The majority of clients identified as Caucasian (60.5%; $n = 1804$), while 11.9% identified as Black or African American ($n = 355$), 8.1% as American Indian or Alaska Native ($n = 241$), 1.3% as Asian ($n = 40$), 0.4% as Native Hawaiian or Pacific Islander ($n = 12$), and 9.7% as multiracial ($n = 288$). Most clients primarily resided with at least one biological or adoptive parent (55.8%; $n = 1664$), while 16.1% of clients lived in a residential treatment center ($n = 480$), 10.5% in kinship care ($n = 312$), 9.5% in foster care (i.e., traditional or treatment; $n = 284$), and 0.8% in a correctional facility ($n = 23$). Approximately 7.2% of clients were identified as residing in another setting (e.g., on their own, homeless, shelter; $n = 216$).

Measures

Clinician measures. Clinicians completed a brief intake questionnaire before participating in the learning collaborative. Demographic data included clinician gender, professional degree, and licensure type.

Fidelity dashboard. Clinician-reported fidelity data were collected at the completion or termination of each TF-CBT case. After starting TF-CBT with a new client, clinicians entered fidelity data into an electronic fidelity “dashboard” (i.e., preformatted Excel spreadsheet). The fidelity dashboard was adapted from the TF-CBT Fidelity Checklist, developed by Deblinger and colleagues (2007). Clinicians added data to the dashboard after each therapy session by specifying the number of minutes spent on

each TF-CBT component: Psychoeducation, Parenting Skills, Relaxation, Affect Expression and Regulation, Cognitive Coping, Trauma Narrative, In-Vivo Desensitization, Conjoint Parent-Child Sessions, Enhancing Safety, and Enhancing Problem Solving. Clinicians could also specify the amount of time they spent implementing material that was not part of the TF-CBT model (i.e., “Other”). When clinicians noted the use of “other” content, they were asked to briefly describe the content covered or technique used. Clinicians also specified the individual or individuals who participated in each component of the TF-CBT session, including the child, caregiver, or both the child and caregiver.

Implementation fidelity composite. TF-CBT implementation fidelity was operationalized by calculating a composite score for each case (i.e., client-clinician pair). The composite score was mirrored after implementation fidelity quantification conducted by Cohen et al. (2016), who derived level of implementation fidelity from treatment length (i.e., duration) and treatment component order (i.e., content). To assess TF-CBT fidelity in a manner more consistent with the implementation fidelity literature (Carroll et al., 2007; Fixsen et al., 2005), the author examined four domains of fidelity: content (i.e., core components), frequency (i.e., rate of occurrence), duration (i.e., length of intervention delivery), and coverage (i.e., clients receiving benefits of intervention) (Carroll et al., 2007). A composite score from 0-8 was derived for each client-clinician pair with points allotted for content (0 = 0 components completed, 1 = 1-2 components completed, 2 = 3-4 components completed, 3 = 5-6 components completed, 4 = 7-8 components completed), frequency of sessions (0 = 0%-74% sessions held at least once per two weeks, 1 = 75%-100% sessions held at least once per two weeks), duration of

treatment (0 = 0-7 sessions, 1 = 31+ sessions, 2 = 8-30 sessions), and coverage (0 = TF-CBT not completed as planned, 1 = TF-CBT completed as planned). Component completion was determined by the number of minutes spent on each component, with 45 minutes deemed the minimum amount of time for component introduction and mastery.

Client measures.

Enrollment questionnaire. Demographic data included client age, gender (Male, Female, Transgender Male/Transgender Female, Other), race (American Indian or Alaska Native, Asian, Black/African American, Native Hawaiian or Other Pacific Islander, White, Unknown), ethnicity (Hispanic or Latino, Not Hispanic or Latino, Unknown), and current primary residence (home with parent(s), residential treatment center, kinship care [with relatives or other family], foster care [treatment and traditional foster care], correctional facility, other [independent, homeless, shelter, other]).

Assessments of youth functioning and trauma exposure. Data were collected using two to four psychometrically valid assessments per client depending on client age. Assessment measures were administered as face-to-face interviews by clinicians, although the *Strengths and Difficulties Questionnaire* could be administered via paper and pencil reporting. Assessments were administered and collected at baseline (i.e., pre-treatment), every three months, and immediately at post-treatment. For the purpose of this study and to remain consistent with the TF-CBT implementation literature, baseline and closing assessments were utilized; if a closing assessment was not completed, the last follow-up assessment was carried forward as a closing assessment (Lang et al., 2015). Clinicians sent the completed assessments to Ambit Network staff via fax or email.

North Shore Trauma History Checklist (NTHC). The NTHC is a 30-item questionnaire that assesses for 17 types of trauma exposure through “Yes/No” responses to questions (North Shore-Long Island Jewish Health System, Inc., 2006). Respondents are also asked about the age of trauma onset, as well as the duration of the trauma exposure. The NTHC is completed by the interviewing clinician, and includes reports from both the child, caregiver, and/or other collateral sources. Trauma types include: disaster, terrorism/political violence, war, serious accident injury, illness/medical trauma, physical assault, physical maltreatment/abuse, neglect, separation from caregiver, homelessness, domestic violence, extreme interpersonal violence, bereavement, community violence, kidnapping/abduction, sexual maltreatment/abuse, sexual assault/rape, and other trauma not reported elsewhere. The last question asks the respondent to determine the “most significant traumatic event(s)” with the following prompt: “Of the things we’ve talked about, which is the worst? Which still really bothers you?”

General Trauma Information Form (GTF). The GTF assesses 22 different types of trauma exposure, in addition to salient characteristics of each exposure (e.g., age, frequency, confirmed or suspected exposure; National Child Traumatic Stress Network, 2014). The GTF is completed by the interviewing clinician, and includes reports from both the child, caregiver, and/or other collateral sources. Trauma types include: (1) Sexual maltreatment/abuse: actual or attempted sexual molestation, exploitation or coercion by a caregiver; (2) Sexual assault/rape: actual or attempted sexual molestation, or coercion not by a caregiver and not recorded as sexual abuse; (3) Trafficking: child has been offered, obtained, or engaged in any kind of trafficking in exchange for anything of

value, money, goods, personal benefit, in-kind favors, or some other kind of benefit; (4) Physical maltreatment/abuse: actual or attempted infliction of physical pain or bodily injury by a caregiver; (5) Physical assault: actual or attempted infliction of physical pain or bodily injury not by a caregiver and not recorded as physical abuse; (6) Psychological or emotional maltreatment/abuse: verbal abuse, excessive demands, emotional neglect; (7) Neglect: physical, medical, or educational neglect; (8) Domestic violence: exposure to physical, sexual, and/or emotional abuse directed at adult caregiver(s) in the home; (9) War/terrorism/political violence inside the U.S.: exposure to any of these events inside the U.S.; (10) War/terrorism/political violence outside the U.S.: exposure to any of these events outside the U.S.; (11) Illness/medical trauma: chronic, life-threatening, or extremely painful illness or medical procedure; (12) Serious accident/injury: unintentional accident or injury; (13) Disaster: major accident or disaster that is the result of a natural or manmade event; (14) Kidnapping/abduction: unlawful seizure or detention against the child's will; (15) Bereavement: death of a loved one, primary caregiver, or sibling; the unexpected, or premature death of a close relative or close friend; (16) Separation: separation from a primary caregiver or sibling; (17) Forced displacement: forced relocation to a different country due to political reasons; (18) Impaired caregiver: history of exposure to caretaker depression/psychiatric problems, other medical illness, or alcohol/drug abuse; (19) Extreme interpersonal violence not reported elsewhere: for example, homicide or suicide; (20) Community violence not reported elsewhere: for example, gang-related violence, neighborhood violence; (21) School violence not reported elsewhere: for example, school shooting, classmate suicide; (22) Bullying: unwanted aggressive behavior among school aged children/peers that involves a real or

perceived power imbalance; (23) Other trauma not reported elsewhere. Additionally, the clinician is asked to specify the primary focus of the current treatment using the 23 aforementioned traumas.

In accordance with definitions established by the National Child Abuse and Neglect Data System (U. S. Department of Health and Human Services, 2011) and the National Child Traumatic Stress Network (Pynoos et al., 2008), a summary variable was created for type of trauma exposure (0 = *non-caregiver-related interpersonal trauma*, 1 = *caregiver-related interpersonal trauma*). Interpersonal trauma was operationalized as exposure to trauma that is of a relational nature by a caregiver, including sexual maltreatment/abuse, physical maltreatment/abuse, psychological or emotional maltreatment/abuse, neglect, domestic violence, separation, and impaired caregiving. The operationalization of caregiver-related interpersonal trauma is also consistent with definitions proffered by D'Andrea and colleagues (D'Andrea, Ford, Stolbach, Spinazzola, & van der Kolk, 2012),

UCLA Posttraumatic Stress Disorder Reaction Index for DSM-IV (UCLA PTSD-RI-IV). The UCLA PTSD-RI-IV is a 48-item questionnaire that assesses for trauma exposure and symptoms of PTSD in children and adolescents (Steinberg et al., 2004; Pynoos et al., 1998), and demonstrates robust reliability and validity (see Elhai et al., 2013; Steinberg et al., 2004; Steinberg et al., 2013). In this sample, Cronbach's alpha was .619 at baseline. Responses are on a 5-point Likert scale, with higher ratings indicating more frequent PTSD symptoms (i.e., 0 = *none of the time*, 4 = *most of the time*). Only the 22-item symptom scales were used in this study. Specifically, 20 items assess PTSD symptoms, while two additional items screen for associated features: fear of recurrence,

and trauma-related guilt. Overall, the UCLA PTSD-RI-IV provides an overall PTSD severity score, and separate scores for each PTSD criterion: Criterion B (Re-Experiencing), Criterion C (Avoidance/Numbing), and Criterion D (Arousal).

UCLA Posttraumatic Stress Disorder Reaction Index for DSM-5 (UCLA PTSD-RI-5). The UCLA PTSD-RI-5 is a 31-item questionnaire that assesses for symptoms of PTSD in children and adolescents (Pynoos & Steinberg, 2013). Although it includes sections on trauma history and exposure details, only the symptom scales were used in this study. Similar to the UCLA PTSD-RI-IV, responses are on a 5-point Likert scale, with higher ratings indicating more frequent symptoms (i.e., 0 = *none of the time*, 4 = *most of the time*). Specifically, 27 items assess PTSD symptoms, while four additional items screen for the dissociative subtype. Overall, the UCLA PTSD-RI-5 provides an overall severity score, and separate scores for each PTSD symptom category: Category B (Intrusion Symptoms), Category C (Avoidance of Stimuli), Category D (Negative Alterations in Cognitions and Mood), and Category E (Marked Alterations in Arousal and Reactivity). Reliability and validity have not yet been published for this measure; however, in this sample, Cronbach's alpha was .770 at baseline.

Strengths and Difficulties Questionnaire—Parent and Self-Report Versions (SDQ). The SDQ is a brief, 25-question behavioral screening measure for children and adolescents with good reliability (Goodman, 2001). Three versions of the SDQ were made available to clinicians, including the Parent and Teacher Versions, which were used with children 3 to 16 years of age, and the Self-Report Version, which was used with children 11 to 16 years of age. Each version has five clinical scales with five items each: Emotional Symptoms, Conduct Problems, Hyperactivity/Inattention, Peer Relationship

Problems, and Prosocial Behavior, and Impact. Each item is rated on a 3-point Likert scale with higher ratings indicating the extent to which each attribute applies to the child (i.e., 0 = *not true*, 2 = *certainly true*).

Data Analysis

Analyses were conducted using SPSS 24.0 (IBM Corp., 2016), including descriptive analyses, common item equating, and multilevel modeling. Distribution of dependent variable normality was checked using measures of skewness, kurtosis, and histogram assessment. Skewness and kurtosis values were within ± 2 and histograms showed normal distributions (Curran et al., 1996). SDQ Self-Report Prosocial Relations showed negative skewness ($S = -.855$, $SE = .171$), while PTSD-RI follow-up ($S = 0.866$, $SE = .100$) showed positive skewness, so data were transformed and re-analyzed. The results were similar to those using raw data. Therefore, in this study, results using raw data are presented. Effect sizes were calculated using an online effect size calculator (Ellis, 2009), and interpreted using Cohen's (1988) original magnitude estimation for Cohen's d : $0.20 = small$, $0.50 = medium$, $0.80 = large$. For analyses investigating factors associated with total implementation fidelity score and implementation fidelity components, t -tests and logistic regression were used to examine associations. The Bonferroni correction for multiple tests of significance was applied and is reported in the text (i.e., $.05/113 = .0004$).

Common item equating. In 2013, the American Psychiatric Association (APA) published new diagnostic criteria for PTSD (APA, 2013) resulting in the 2015 replacement of the UCLA PTSD-RI-IV (Pynoos et al., 1998) with the UCLA PTSD-RI-5 (Pynoos & Steinberg, 2013). In line with this change, the Ambit Network adopted the

PTSD-RI-5 in 2015 after using the UCLA PTSD-RI-IV from 2007 to 2015.

Approximately 80% (n baseline= 1,967) of the Ambit Network's UCLA PTSD-RI data were in the form of the UCLA PTSD-RI-IV, while 20% (n baseline = 467) were in the form of the UCLA PTSD-RI-5. In order to use data from both measures, common item equating was used to equate both measures resulting in the "PTSD-RI" dependent variable. In common item equating, Kolen and Brennan (2004) posit that a common item set should be at least 20% of the length of a total test containing 40 or more items.

Although the two versions of the UCLA PTSD-RI are shorter than 40 items, 86.4% (n = 19 of 22) of the UCLA PTSD-RI-IV items and 70.4% (n = 19 of 27) of the UCLA PTSD-RI-5 items overlap. These proportions exceed Kolen and Brennan's (2004) common item set criterion, suggesting little random equating error.

Common items from the UCLA PTSD-RI-IV and 5 were determined through an item-by-item analysis of both assessment measures. Specifically, the UCLA PTSD-RI-IV is composed of 22 items across three symptom clusters (i.e., B: Re-Experiencing, C: Avoidance/Numbing, D: Increased Arousal), while the UCLA PTSD-RI-5 is comprised of 27 items across four symptom clusters (i.e., B: Intrusion, C: Avoidance, D: Negative Alterations in Cognition and Mood, and E: Alterations in Arousal and Reactivity). Nineteen items were found to link across UCLA PTSD-RI versions, resulting in a range of 0 to 76 points. Missing PTSD-RI common item data were calculated using mean substitution. Nineteen items were found to link across UCLA PTSD-RI versions, resulting in a range of 0 to 76 points. Missing PTSD-RI common item data were calculated using mean substitution. At baseline, Cronbach's alphas for the PTSD-RI were .896 for *DSM-IV* and .904 for *DSM-5* common items.

Multilevel modeling. Given the nested nature of the data, multilevel modeling was used to account for the dependency that may result from clustered or non-independent data (Raudenbush & Bryk, 2002). Specifically, multilevel modeling produces unbiased standard error estimates associated with regression coefficients and allows for group characteristics to be included in statistical models of individual outcomes (Raudenbush & Bryk, 2002). The hypothesized model for this study consisted of four levels: timepoints (level 1) were nested within clients (level 2: $n = 2,980$), who were nested within clinicians (level 3: $n = 596$), and clinicians were nested within agencies (level 4: $n = 75$).

Intraclass correlation coefficients. Intraclass correlation coefficients (ICCs) were calculated to determine the proportion of variance in the dependent variables explained by linear mixed model grouping structures. ICCs were calculated using unconditional linear mixed modeling for dependent variable delta scores (i.e., follow-up – baseline) to assess for 2-level (i.e., client-clinician, client-agency) and 3-level (i.e., client-clinician-agency) clustering of data. Both 2- and 3-level models are referred to as *clinician* (time-client-clinician) and *agency* (time-client-agency) models for ease of reference. Nezlek (2008) estimated that ICCs close to zero are low, while ICCs close to one are high. ICCs for outcome delta scores are presented in Table 13. When calculating ICCs for the SDQ Self-Report Total and SDQ Parent-Report Total scale deltas, all 3-level models failed to converge. ICCs for PTSD-RI were 0.03 for the 2-level clinician model, 0.009 for the 2-level agency model, and 0.007 for the 3-level clinician and 0.008 for the 3-level agency model. Across all variables, results suggested that multilevel modeling for outcome change analyses was not required. However, in order to perform the most conservative

analyses possible, multilevel modeling was utilized to account for marginal/minimal clustering and missing data.

Applying multilevel models. First, 3-level mixed models were applied (i.e., random intercept and slope, random intercept only, and random slope only models) and converged for SDQ Self-Report Peer Problems, SDQ Self-Report Impact, SDQ Parent Report Total, SDQ Parent Report Impact, and PTSD-RI delta scores. Next, 2-level mixed models were applied for all scale deltas. Random intercept and slope models failed to converge, thus random intercept only and random slope only models were tested. Although all convergence criteria were satisfied for random slope models, the final Hessian matrix was only positive for SDQ Self-Report Hyperarousal-Inattention (clinician), SDQ Parent Report Hyperarousal-Inattention (agency), and SDQ Parent Report Prosocial Relations (agency) scales. Thus, results from the 2-level random slope models are reported for those variables. For SDQ Self-Report Total (clinician), SDQ Self-Report Conduct Problems (clinician), SDQ Self-Report Prosocial Relations (agency), SDQ Parent Report Conduct Problems (agency), SDQ Parent-Report Emotional Symptoms (agency), and SDQ Parent-Report Peer Problems (clinician), the random intercept model converged and results from the analyses are reported. For scales in which multiple models converged, final models were selected by comparing Akaike information criterion (AIC) and Bayes information criterion (BIC) values (i.e., smaller values indicated better fit), which yielded random intercept only models for SDQ Self-Report Peer Problems (clinician) SDQ Self-Report Impact (agency), SDQ Parent Report Total (clinician), SDQ Parent Report Impact (agency), and PTSD-RI (agency). Two-level

random intercept only and random slope only models failed to converge for SDQ Self-Report Emotional Symptoms necessitating analysis of variance (ANOVA).

For all models, the association between follow-up and baseline delta scores and implementation fidelity to the TF-CBT model was the a priori comparison (Delta). All models controlled for the number of months between baseline and follow-up assessments (Months in Treatment). When fitting linear mixed models, SPSS uses the Satterthwaite approximation method to compute denominator degrees of freedom. Thus, denominator degrees of freedom used for *F* statistics are not whole numbers (West, 2009).

Missing data. Missing values were present in the dataset due to the longitudinal nature of the research design. Missing data in all models were managed using the restricted maximum likelihood (REML) procedure available through SPSS 24.0.

Results

TF-CBT Implementation Outcomes

Learning collaborative completion. Approximately 90% of clinicians ($n = 535$; 89.8%) who began training in TF-CBT completed the learning collaborative model. Sixty-one clinicians left their learning collaboratives for a variety of reasons: left agency ($n = 18$; 29.5%), removed from cohort due to concerns ($n = 12$; 19.7%), illness ($n = 1$; 1.6%), or an unknown reason ($n = 30$; 5.0%).

Implementation fidelity. Complete implementation fidelity data were available for 629 clients (i.e., 21.1% of the study sample). Implementation fidelity data are presented in Table 14. The average implementation fidelity composite score was 4.7 ($SD = 2.31$) and ranged from 0 to 8. Regarding the content of the TF-CBT model implemented, over half of clients who began TF-CBT received Psychoeducation ($n =$

512; 81.4%), Affective Expression and Regulation ($n = 444$; 70.6%), Relaxation ($n = 401$; 63.8%), Trauma Narration and Processing ($n = 355$; 56.4%), and Cognitive Coping ($n = 346$; 55.0%). Less than half of clients and their families received Parenting Skills ($n = 241$; 38.3%), Safety Enhancement ($n = 21.9\%$), and Conjoint Witnessing ($n = 127$; 20.9%). In terms of the rate of session occurrence, 64.4% of clients attended weekly or biweekly sessions at least 75% of the time ($n = 405$), with the majority of clients participating in 8 to 30 sessions ($n = 377$; 59.9%). TF-CBT coverage was calculated for the full sample of 2,980 clients. Overall, 20.7% of the sample completed TF-CBT as planned ($n = 616$), while 20.4% dropped out of TF-CBT treatment ($n = 607$). Moreover, clinicians indicated that 6.3% of youth were lost to follow up ($n = 187$) and 5.4% were transferred to another TF-CBT-trained clinician due to youth mobility ($n = 160$). Most clients discontinued TF-CBT for an unspecified reason ($n = 1,410$; 47.3%).

Association between TF-CBT Implementation Fidelity and Change in Client

Outcomes

Table 15 presents baseline and follow-up means, change scores, and effect sizes for client outcomes.

Correlations between implementation fidelity and change in client outcome variables. Correlations between the implementation fidelity composite and client outcomes are presented in Table 16. Regarding total implementation fidelity score, implementation fidelity was positively associated with SDQ Parent-Reported Prosocial Relations ($r = 0.26, p < .01$). When examining associations with implementation fidelity components, frequency was statistically significantly related to improvement in SDQ Parent-Reported Conduct Problems ($r = 0.34, p < .01$); while coverage was statistically

significantly related to improvements in SDQ Self-Reported ($r = 0.18, p < .05$) and SDQ Parent-Reported Emotional Symptoms ($r = 0.14, p < .05$).

Multilevel modeling of association between implementation fidelity and change in client outcomes.

PTSD-RI. Table 17 presents multilevel modeling results for the association between implementation fidelity and PTSD-RI change scores. Holding months in treatment constant, results revealed that for every 1-point increase in implementation fidelity composite score, there was a 0.48 increase in PTSD-RI change score (delta), $F(1, 134.85)=0.81, p = .370$. Although non-significant, this result indicated that an increase in implementation fidelity score is related to greater improvements in PTSD symptomatology. Estimates of covariance parameters demonstrated trend-level variability in the mean scores (i.e., intercepts) between the PTSD-RI change score and implementation fidelity across agencies, Wald $Z = 1.77, p = 0.077$.

SDQ Self-Report. Table 18 presents multilevel modeling results for the association between SDQ Self-Report scale change scores and implementation fidelity composite score. Holding months in treatment constant, results suggested that for every 1-point increase in implementation fidelity composite score, there was a 0.003 increase in the change score for the SDQ Self-Report Total, $F(1, 43.13)=0.00, p = .993$; 0.07 increase in the change score for SDQ Self-Report Emotional Symptoms, $F(2, 56)=0.25, p = .666$; 0.07 increase in change score for SDQ Self-Report Conduct Problems, $F(1, 49.17)=0.42, p = 0.52$; .05 increase in change score for SDQ Self-Report Hyperactivity-Inattention, $F(1, 43.44)=0.17, p = .686$; -0.07 increase in change score for SDQ Self-Report Prosocial Relations, $F(1, 53.99)=0.29, p .594$; and a 0.06 increase in the change score for SDQ

Self-Report Impact, $F(1, 45.29)=0.14, p = .707$. Contrary to expectations, holding months in treatment constant, results suggested that for every 1-point increase in implementation fidelity composite score, there was a 0.02 decrease in the change score for SDQ Self-Report Peer Problems, $F(1, 21.87)=0.06, p = .811$. Estimates of covariance parameters demonstrated significant variability in the mean scores (i.e., intercepts) between the SDQ Self-Report Peer Problems change score and implementation fidelity composite score across clinicians, Wald $Z = 3.31, p = 0.001$.

SDQ Parent Report. Table 19 presents the multilevel modeling results for the association between SDQ Parent Report scale change scores and implementation fidelity composite score. Holding months in treatment constant, results suggested that for every 1-point increase in implementation fidelity composite score, there was a 0.10 increase in the change score for SDQ Parent Report Conduct Problems, $F(1,53.20)=0.39, p = .534$; -0.25 increase in change score for SDQ Parent Report Prosocial Relations, $F(1, 52.82)=3.68, p = 0.06$; and 0.06 increase in change score for SDQ Parent Report Impact score, $F(1, 45.29)=0.14, p = .707$. Contrary to expectations, holding months in treatment constant, results suggested that for every 1-point increase in implementation fidelity composite score, there was a 0.39 decrease in the change score for SDQ Parent Report Total, $F(1, 47.67)=0.87, p = .356$; 0.06 decrease in the change score for SDQ Parent Report Hyperactivity-Inattention, $F(1, 52.70)=0.12, p = .731$; and 0.23 decrease in change score for SDQ Parent Report Peer Problems, $F(1, 53.97)=2.82, p = .099$. Trend-level (i.e., $p < .10$) associations emerged for the relationship between implementation fidelity composite score and SDQ Parent Report Peer Problems as well as Prosocial Relations change scores: increases in implementation fidelity composite score were

related to trend-level greater improvements on SDQ Parent Report Prosocial Relations and worsening on Peer Problems scale scores. Estimates of covariance parameters demonstrated no significant variability in the mean scores (i.e., intercepts) or rates of change (i.e., slopes) between SDQ Parent Report change scores across clinicians or agencies, respectively.

Factors Associated with TF-CBT Implementation Fidelity

Agency-, clinician-, and client-level factors associated with TF-CBT implementation fidelity composite scores are shown in Table 20.

Agency-level factors. Correlational analyses revealed a trend-level association between type of agency and total implementation fidelity score ($r = 0.08, p < .10$), and no association between number of sites and total implementation fidelity score. Follow-up t -tests revealed that clinicians from agencies providing both outpatient and in-home services ($M = 4.51, SD = 2.38$) demonstrated statistically significantly lower implementation fidelity scores than clinicians from outpatient-only agencies ($M = 5.34, SD = 2.06$), $t(35.16) = 2.05, p = .047$; and comprehensive community agencies ($M = 5.55, SD = 2.20$), $t(240) = -1.98, p = .049$.

Regarding components of implementation fidelity, a statistically significant association emerged between duration of TF-CBT and type of agency ($r = 0.10, p < .05$). Follow-up t -tests indicated that clinicians from comprehensive community agencies ($M = 1.86, SD = 0.47$) implemented TF-CBT for a more preferred duration (i.e., 8-30 sessions) than clinicians from outpatient-community ($M = 1.24, SD = 0.91$), $t(33.13) = -5.57, p < .001$; outpatient-intensive ($M = 1.44, SD = 0.87$), $t(68.47) = -2.82, p = .006$; outpatient-

community-intensive ($M = 1.30$, $SD = 0.85$), $t(59.81) = -4.23$, $p < .001$; and intensive ($M = 1.38$, $SD = 0.82$), $t(58.23) = -3.64$, $p = .001$) agencies.

Clinician-level factors.

Gender. Clinician gender was statistically significantly related to the frequency of TF-CBT sessions ($r = 0.10$, $p < .05$). Logistic regression indicated that the odds of implementing 75 to 100 percent of sessions at least once every two weeks were 1.78 times higher for female clinicians than male clinicians, $\chi^2(1) = 5.77$, $OR = 1.78$, $p = 0.016$.

License. Clinician license was statistically significantly associated with total implementation fidelity score. ($r = 0.09$, $p < .05$). Follow-up t -tests indicated that clinicians with licenses in professional counseling ($M = 5.23$, $SD = 2.05$) demonstrated statistically significantly higher implementation fidelity scores than clinicians with social work licenses ($M = 4.55$, $SD = 2.35$), $t(196.85) = -2.81$, $p = .005$; and marriage and family therapy licenses ($M = 4.36$, $SD = 2.52$), $t(161.28) = -2.56$, $p = .01$.

Regarding implementation fidelity components, statistically significant associations emerged for clinician license and content ($r = 0.09$, $p < .05$), as well as duration ($r = 0.12$, $p < .01$). With respect to TF-CBT content, follow-up t -tests indicated that clinicians with licenses in professional counseling ($M = 2.45$, $SD = 1.15$) and psychology ($M = 2.43$, $SD = 1.01$) completed more TF-CBT components than clinicians with marriage and family therapy licenses ($M = 2.13$, $SD = 1.24$), $t(184) = -1.80$, $p = .074$ and $t(166) = -1.74$, $p = .084$, respectively. However, these differences were only significant at a trend-level. Regarding duration of TF-CBT implementation, t -tests revealed that clinicians with licenses in professional counseling ($M = 1.48$, $SD = 0.76$) and psychology ($M = 1.47$, $SD = 0.83$) implemented TF-CBT for a more preferred duration

than clinicians with social work licenses ($M = 1.24$, $SD = 0.90$), $t(204.32) = -2.61$, $p = .010$ and $t(140.78) = -2.23$, $p = .028$, respectively.

Supervisory status. Supervisory status was statistically significantly associated with total implementation fidelity score ($r = 0.08$, $p < .05$). Follow-up t -tests revealed that clinicians who provided supervisory services demonstrated statistically significantly higher implementation fidelity scores ($M = 5.14$, $SD = 2.16$) than non-supervisors ($M = 4.63$, $SD = 2.33$), $t(627) = -2.02$, $p = .044$.

Regarding components of implementation fidelity, statistically significant associations emerged for supervisory status and content ($r = 0.09$, $p < .05$), as well as coverage ($r = 0.08$, $p < .01$). A follow-up t -test revealed that supervisors completed statistically significantly more TF-CBT components ($M = 2.55$, $SD = 0.11$) than non-supervisors ($M = 2.24$, $SD = 1.18$), $t(627) = -2.30$, $p = .022$. Additionally, logistic regression indicated that the odds of completing TF-CBT were 1.60 times higher for supervisors than non-supervisors, $\chi^2(1) = 17.51$, $OR = 1.60$, $p < .001$. This test withstood the Bonferroni correction.

Client-level factors. No client-level factors were statistically significantly related to implementation fidelity composite score, although there was a trend-level finding for exposure to interpersonal trauma ($r = -0.07$, $p < .10$). Follow-up analyses revealed that total implementation fidelity scores were higher for clients exposed to non-interpersonal traumatic events ($M = 5.16$, $SD = 2.37$) than clients exposed to interpersonal traumatic events ($M = 4.65$, $SD = 2.30$), $t(627) = 1.71$, $p = .088$.

Age. Client age was statistically significantly related to TF-CBT coverage ($r = -0.06$, $p < .01$). Logistic regression revealed that for every one-year increase in age, the

relative rate of TF-CBT completion was 0.96 or 4 percent lower, $\chi^2(1)=17.51$, $RR=0.96$, $p = .003$.

Ethnicity. Client ethnicity was statistically significantly related to TF-CBT duration ($r = 0.12$, $p < .01$). Clients who did not identify a specific ethnicity ($M = 1.72$, $SD = 0.62$) participated in TF-CBT for a more preferred duration than clients who identified as Hispanic/Latino ($M = 1.21$, $SD = 0.89$), $t(116.53)=-3.74$, $p < .001$ and not Hispanic/Latino ($M = 1.31$, $SD = 0.88$), $t(64.93)=-4.23$, $p < .001$. These tests withstood the Bonferroni correction.

Primary residence. Primary residence was statistically significantly related to coverage ($r = -0.06$, $p < .01$). Logistic regression indicated that the odds of completing TF-CBT were 1.60 times higher for clients residing in residential treatment centers compared with clients living elsewhere, $\chi^2(1)=17.03$, $OR=1.60$, $p < .001$; and 1.42 times higher in clients residing in traditional foster care compared to clients living elsewhere, $\chi^2(1)=5.35$, $OR=1.42$, $p = .021$. Finally, the odds of completing TF-CBT were 1.36 times higher for clients living outside the home than clients living at home with biological or adoptive parents, $\chi^2(1)=11.37$, $OR=1.36$, $p = .001$.

Total number of traumatic events. The total number of traumatic events was statistically significantly associated with coverage ($r = 0.06$, $p < .01$). Logistic regression indicated that for every additional trauma exposure, the relative rate of TF-CBT completion was 1.06 or 6 percent higher, $\chi^2(1)=12.11$, $RR=1.06$, $p = .001$.

Interpersonal trauma exposure. Interpersonal trauma exposure was statistically significantly associated with TF-CBT content ($r = -0.08$, $p < .05$) and coverage ($r = 0.07$, $p < .01$). Regarding content of TF-CBT, clients exposed to interpersonal trauma ($M =$

2.26, SD = 1.17) completed statistically significantly fewer TF-CBT components than clients who had not experienced interpersonal trauma (M = 2.56, SD = 1.16), $t(627)=2.00, p = .046$. However, logistic regression revealed that the odds of completing TF-CBT were 1.55 times higher for clients exposed to interpersonal trauma than clients who did not report exposure to interpersonal trauma, $\chi^2(1)=12.50, OR=1.55, p < .001$. This test withstood the Bonferroni correction.

Discussion

Despite the increasing number of trauma-focused EBTs for youth and their families, there is an urgent need to investigate the transportability to routine community practice settings on a scale large enough to improve the overall mental health of trauma-exposed populations. The current study investigated the nine-year statewide implementation of TF-CBT by the University of Minnesota's Ambit Network, and is the first TF-CBT implementation initiative to examine the relationship between clinician-reported TF-CBT implementation fidelity and client outcomes in routine community practice settings. The author also examined agency-, clinician-, and client-level factors with potential to impact TF-CBT implementation fidelity and its components (i.e., content, coverage, duration, frequency) in the context of a large-scale implementation initiative.

Multilevel modeling indicated that implementation fidelity scores were not significantly related to change in posttraumatic stress symptomatology from baseline to follow-up among trauma-exposed youth. Although the relationship was not significant, associations were in the anticipated direction such that increases in implementation fidelity scores were related to greater improvements in PTSD symptomatology. The

association between implementation fidelity and most SDQ Self-Report change scores were similarly non-significant, although in the anticipated direction for change in total score, emotional symptoms, conduct problems, hyperactivity-inattention, prosocial relations, and overall impact of difficulties. Contrary to expectations, the association between self-reported peer problems and implementation fidelity was negative suggesting that greater fidelity scores were related to worsening of peer problems, albeit non-significant. Finally, multilevel findings from SDQ Parent Report change scores revealed a trend-level positive relationship between implementation fidelity and change in prosocial relations, and a negative relationship between implementation fidelity and change in peer problems. All other associations between SDQ Parent Report change scores and TF-CBT were not significant, but positive for prosocial relations and overall impact of difficulties, and negative for total score, and hyperactivity-inattention.

It is important to recognize the differences in implementation fidelity and child outcome change score relationship between client and caregiver report, in that better TF-CBT implementation fidelity was related to worsening of total emotional and behavioral problems, peer problems, and hyperactivity-inattention. The caregiver-report findings are surprising given previous research on the impact of implementation fidelity on client outcomes. For example, a 2008 review of five meta-analyses investigated the impact of various prevention and intervention programs (e.g., anti-bullying programs, drug use prevention studies, school-based prevention programs targeting aggressive behaviors) on client outcomes (Durlak & DuPre, 2008). Results revealed that interventions with better, as opposed to poorer, implementation fidelity demonstrated effect sizes that were two to three times larger. Some research, however, has proffered contradictory evidence

complicating this trend. Barber and colleagues (2006), for instance, found that “perfect” adherence to a drug counseling protocol was less predictive of positive intervention outcomes than a moderate level of adherence (Barber, Sharpless, Lkostermann, & McCarthy, 2007). It may be that rigid adherence to intervention protocols reduces clinician flexibility and adaptability that are needed to tailor interventions to local client populations (Breitenstein et al., 2010).

The literature on child-caregiver symptom report agreement provides a second interpretation in that a number of studies have found that caregivers tend to underestimate their children’s PTSD symptomatology, particularly externalizing posttraumatic stress symptomatology (Ceballo et al., 2001; Shemesh et al., 2005; Stover et al., 2010).

Therefore, it may be possible that clinicians with better implementation fidelity more readily included caregivers in the TF-CBT treatment model and provided better psychoeducation about the effects of trauma and symptoms of PTSD. Improved understanding and recognition of trauma sequelae may have resulted in increased insight into—or overestimation of—children’s internalizing and externalizing symptoms.

Results from the present study also revealed specific agency-, clinician-, and client-level characteristics that were associated with better implementation fidelity to the TF-CBT model. Regarding agency-level factors, there was a trend-level association between the type of agency and overall implementation score, with clinicians from outpatient-only and comprehensive community agencies implementing TF-CBT with greater fidelity. Clinicians from comprehensive community agencies also implemented TF-CBT for a more preferred duration of time (i.e., 8-30 sessions) compared to clinicians providing services in other community practice settings.

Although agency-level factors have not been examined in TF-CBT implementation fidelity research, Aarons and colleagues (Aarons et al., 2011) suggest that agencies with greater absorptive capacity—or the preexisting knowledge, ability to use knowledge, and specialization—are more likely to explore and implement EBTs. Public sector agencies, such as child welfare or juvenile justice, have been shown to demonstrate less absorptive capacity due to workforces with varied levels of education, heavy caseloads, and responsibility to deliver varied services (Yoo et al., 2007). Within the state of Minnesota, it may be that outpatient-only and comprehensive community agencies have greater absorptive capacity to adopt and sustain TF-CBT, although more research is needed to replicate these findings.

Clinician-level characteristics were also variably related to implementation fidelity, with clinician license and supervisory status statistically significantly related to implementation fidelity score. Clinicians with professional counseling licenses (e.g., Licensed Professional Counselors, Licensed Professional Clinical Counselors) provided TF-CBT with higher overall levels of implementation fidelity, completed more TF-CBT components, and provided TF-CBT for a more preferred duration of time (i.e., 8 to 30 sessions). Clinicians with psychology licenses were also found to complete more TF-CBT components and provide TF-CBT for a more preferred duration of time. These findings contrast with previous examinations of clinician-level outcomes and their relationship with TF-CBT implementation fidelity. Allen and Johnson (2012), for example, surveyed 132 mental health clinicians in children's advocacy centers to assess the likelihood of implementation of TF-CBT model components. Analyses suggested that the likelihood of regularly implementing all components was not predicted by

professional discipline (i.e., counseling, social work, clinical psychology) or years of experience. Moreover, in a randomized implementation study of TF-CBT for adjudicated teens in residential treatment centers, clinicians' professional discipline did not significantly impact TF-CBT implementation outcomes (i.e., fidelity composite of duration and content) (Cohen et al., 2016). Of course, it could be argued that clinician license is not commensurate with professional discipline.

Regarding supervisory status results, clinicians who provided supervisory services provided TF-CBT with higher overall levels of implementation fidelity, delivered more TF-CBT components, and completed TF-CBT more often than non-supervisors. Although implementation science has investigated the impact of regular supervision or coaching on implementation fidelity (McHugh & Barlow, 2010), little or no evidence proffers that supervisors provide EBT services with greater fidelity. Better implementation fidelity by supervisors may point to the importance of experience (i.e., inferring that supervisors have more experience) or receipt of additional training or coaching support. During Ambit Network TF-CBT learning collaboratives, for instance, supervisors received additional consultative support in the form of didactic content during in-person training and supervision consultation calls, which may be related to enhanced TF-CBT implementation fidelity.

Finally, no client-level factors were significantly related to total implementation fidelity score; although exposure to interpersonal trauma was associated with lower implementation fidelity scores, this was only observed at a trend level. Regarding TF-CBT coverage, completing TF-CBT as planned occurred more frequently among clients residing outside the home (i.e., particularly in residential treatment centers and traditional

foster care), clients exposed to interpersonal trauma, clients reporting more trauma exposure, and younger clients. It is difficult to reconcile the findings that clients' exposure to interpersonal trauma led clinicians to implement TF-CBT with somewhat less fidelity, but with significantly greater odds of completion. Such discrepancy may lend credence to the idea that implementation fidelity is multifactorial in nature and these factors interact in unpredictable ways. For example, it may be the case that clients exposed to caregiver-related interpersonal trauma or multiple traumatic events are more frequently involved with Child Protective Services (CPS), which may require the client's participation in and completion of TF-CBT after removal from parental custody.

Limitations

The limitations of this study are related to the real-world, or community practice, context in which TF-CBT training and implementation took place. Implementation outcomes (i.e., implementation fidelity) were based on clinician self-reports, which may not be as accurate as observational or tape ratings (Carroll, Nich, & Rounsville, 1998; Henggeler, Melton, Brondino, Scherer, & Hanley, 1997). However, some evidence suggests high inter-rater reliability (.92) on the TF-CBT Fidelity Checklist between community clinicians' self-ratings of the fidelity of their own sessions and expert ratings of the same audiotaped session (Cohen, Mannarino, & Iyengar, 2011). Moreover, although data on TF-CBT implementation fidelity was available for 629 youth-clinician pairs, there may have been an overrepresentation of fidelity data for clinicians providing TF-CBT while participating in the learning collaborative model given that clinicians were required to complete self-reported fidelity measures during training. Generally, in order to sustain implementation fidelity over time, rigorous methods and measures are

generally required to prevent fidelity decay or drift (Forgatch & DeGarmo, 2011; McHugh & Barlow, 2010; Waller, 2009). According to Forgatch and DeGarmo (2011), the continued implementation of EBTs to community practice settings is a dynamic process and requires extensive collaboration and long-term commitment and may be influenced by “poor training, lack of theory, unregulated adaptations, failure to monitor, or some combination of these” (p. 245).

Finally, implementation fidelity research generally is limited by insufficient and poorly validated assessment tools (Klimes-Dougan et al., 2009). In the context of this investigation, the TF-CBT Fidelity Checklist is still in the early stages of development and its validity has yet to be fully established. Although this investigation measured TF-CBT implementation fidelity using elements of the TF-CBT Fidelity Checklist, there is no definitive evidence that the TF-CBT implementation fidelity composite accurately captured clinician behaviors and other processes congruent with TF-CBT’s underlying theoretical framework (Breitenstein et al., 2010).

Conclusion

The present TF-CBT implementation initiative is the first to examine the relationship between TF-CBT implementation fidelity and change in client outcomes within ecologically valid, routine practice settings. Findings from this study also bring attention to multiple agency-, clinician-, and client-level factors that may impact implementation fidelity when scaling up TF-CBT, or other trauma-focused EBTs, in community practice settings. In light of the limited financial and staffing resources available to build capacity to adopt and sustain TF-CBT training and utilization,

continued research in trauma-focused EBT implementation is urgently needed (Bond, Glover, Godfrey, Butler, & Patton, 2001; Klimes-Dougan et al., 2009).

Table 11

Agency Characteristics (N = 75)

	<u>μ</u>	<u>SD</u>
Number of Sites	3.44	3.32
	<u>n</u>	<u>%</u>
1	34	45.3
2-4	16	21.3
5-7	19	25.3
8+	6	8.0
Services Offered		
Outpatient and In-Home	33	44.0
Outpatient	13	17.3
Outpatient, In-Home, and Inpatient	11	14.7
Outpatient and Inpatient	7	9.3
Inpatient	5	6.7
Comprehensive	3	4.0
Correction	2	2.7
In-Home	1	1.3

Note. Comprehensive agencies offered outpatient, in-home, inpatient, and correctional services.

Table 12

Clinician Characteristics (N = 596)

	<u><i>n</i></u>	<u><i>%</i></u>
Gender		
Male	106	17.8
Female	490	82.2
Degree		
Masters of Social Work (MSW)	206	34.6
Masters of Arts (MA)	202	33.9
Masters of Science (MS)	98	16.4
Doctor of Psychology (PsyD)	33	5.5
Doctor of Philosophy (PhD)	30	5.0
Masters of Education	5	0.8
Other or Unknown	22	3.7
Licensure		
Social Work (LICSW, LGSW, LISW)	211	35.4
Marriage and Family Therapy (LMFT, LAMFT)	143	24.0
Professional Counseling (LPC, LPCC)	124	20.8
Psychology	48	8.1
Other or Unknown	70	11.7

Table 13

Intraclass Correlation Coefficients for Client Outcome Deltas

<u>Measure</u>	<u>3-Level (Clinician, Agency)</u>	<u>2-Level (Clinician)</u>	<u>2-Level (Agency)</u>
PTSD-RI	Clinician = 0.007 Agency = 0.008	0.03	0.009
SDQ Self Report	--	--	--
Total	--	--	--
Emotional	--	--	--
Conduct Problems	--	--	--
Hyperarousal-Inattention	--	0.007	--
Peer	--	0.08	0.02
Prosocial	--	--	0.02
Impact	--	0.09	0.02
SDQ Parent Report	--	--	--
Total	--	--	--
Emotional	--	--	0.004
Conduct Problems	--	0.009	--
Hyperarousal-Inattention	--	--	--
Peer	--	--	--
Prosocial	--	0.04	0.02
Impact	--	0.09	--

Note: -- = no convergence

Table 14

Clinician Implementation Fidelity: Composite and Components Scores (n = 629)

	<u>μ</u>	<u>SD</u>
Fidelity Composite (range = 0-8)	4.71	2.31
	<u>n</u>	<u>%</u>
Content		
Psychoeducation	512	81.4
Parenting Skills	241	38.3
Relaxation	401	63.8
Affective Expression/Regulation	444	70.6
Cognitive Coping	346	55.0
Trauma Narrative	355	56.4
Conjoint Witnessing	127	20.2
Enhancing Safety	138	21.9
Frequency		
0%-74%	224	35.6
75%-100%	405	64.4
Duration		
0-7	170	27.0
8-30	377	59.9
31+	82	13.0
Coverage (N=2,980)		
Completed as Planned	616	20.7
Client Dropped Out	607	20.4
Case Lost	187	6.3
Case Transferred	160	5.4
Other	1,410	47.3

Note: Content refers to TF-CBT components completed; frequency refers to percent weekly or biweekly sessions completed; duration refers to the number of sessions completed; and coverage refers to TF-CBT completed as planned.

Table 15

Client Outcomes, Deltas, and Effect Sizes

<u>Measure</u>	<u>n</u> (B/F)	<u>Baseline</u> μ (SD)	<u>Follow-Up</u> μ (SD)	<u>Δ (SD)</u>	<u>Effect Size (<i>d</i>)</u>
PTSD-RI	2,495/621	33.17 (15.87)	20.72 (14.51)	13.65 (15.61)	0.92
SDQ Self-Report					
Total	769/274	16.63 (6.30)	13.92 (6.35)	2.66 (6.14)	0.42
Emotional Symptoms	770/276	4.70 (2.59)	3.81 (2.61)	1.01 (2.42)	0.38
Conduct Problems	770/275	3.28 (2.07)	2.46 (1.96)	0.68 (1.96)	0.33
Hyperactivity- Inattention	770/276	5.51 (2.39)	4.88 (2.41)	0.64 (2.15)	0.27
Peer Problems	770/276	3.23 (2.18)	2.72 (1.84)	0.44 (1.87)	0.23
Prosocial Relations	770/276	7.61 (2.08)	7.88 (1.96)	-0.25 (1.98)	0.12
Impact	713/259	2.52 (2.58)	1.53 (2.14)	0.61 (2.42)	0.27
SDQ Parent Report					
Total	999/309	17.93 (7.39)	14.66 (7.21)	3.63 (6.55)	0.49
Emotional Symptoms	1,001/310	4.75 (2.65)	3.78 (2.56)	1.24 (2.77)	0.46
Conduct Problems	1,002/311	3.91 (2.60)	3.00 (2.42)	0.87 (2.27)	0.35
Hyperactivity- Inattention	1,002/311	5.91 (2.84)	5.24 (2.65)	0.91 (2.47)	0.33
Peer Problems	1,001/311	3.43 (2.23)	2.65 (2.08)	0.59 (1.97)	0.26
Prosocial Relations	1,001/311	6.93 (2.34)	7.50 (2.12)	-0.29 (2.06)	0.15
Impact	939/287	3.76 (2.93)	2.34 (2.63)	0.61 (2.42)	0.52

Note. B: Baseline; F: Last Follow Up or Closing; UCLA PTSD-RI: UCLA

Posttraumatic Stress Disorder Reaction Index; SDQ: Strengths and Difficulties

Questionnaire; d: Cohen's d.

Table 16

Bivariate Correlations (r) between Outcome Deltas and Implementation Fidelity Composite Scores

<u>Measure</u>	<u>Fidelity Components</u>				<u>Fidelity</u>
	<u>Content</u>	<u>Frequency</u>	<u>Duration</u>	<u>Coverage</u>	<u>Total</u>
PTSD-RI	0.01	-0.03	0.07	0.07	0.04
SDQ Self-Report					
Total	-0.04	0.00	0.02	-0.03	-0.00
Emotional Symptoms	0.09	-0.04	-0.01	0.11	-0.07
Conduct Problems	0.01	0.01	0.10	-0.10	0.03
Hyperac.-Inattention	-0.02	0.05	0.11	0.03	0.05
Peer Problems	-0.11	0.03	-0.17	-0.11	-0.08
Prosocial Relations	-0.03	0.05	-0.11	-0.04	-0.06
Impact	-0.04	0.02	0.04	0.18*	0.04
SDQ Parent-Report					
Total	-0.16	0.26 [†]	-0.16	0.12 [†]	-0.05
Emotional Symptoms	-0.09	0.19	-0.04	0.14*	0.01
Conduct Problems	-0.03	0.34**	-0.04	0.10	0.09
Hyperac.-Inattention	-0.10	0.13	-0.14	0.04	-0.07
Peer Problems	-0.23	0.06	-0.23 [†]	0.03	-0.19
Prosocial Relations	0.21	0.06	0.25 [†]	0.11 [†]	0.26**
Impact	-0.23 [†]	0.18	-0.19	0.06	-0.13

** $p < .01$, * $p < .05$, [†] $p < .10$

Table 17

PTSD-RI Deltas: Estimated Fixed Effects and Covariance Parameters

	β	$SE(\beta)$	t	p
Implementation Fidelity	.48	0.53	0.90	.370
Months in Treatment	-.03	0.30	-0.11	.914
Estimates of Covariance				
	<u>Estimate</u>	<u>SE</u>	<u>Wald Z</u>	<u>p</u>
Agency	36.19	20.46	1.77	.077

Table 18

SDQ Self-Report Deltas: Estimated Fixed Effects and Covariance Parameters

Fixed Effects				
	β	SE(β)	t	p
Total Score				
Implementation Fidelity	0.00	0.41	0.01	.993
Months in Treatment	-0.03	0.18	-0.14	.886
Emotional Symptoms				
Implementation Fidelity	0.07	0.16	0.43	.666
Months in Treatment	0.04	0.08	0.53	.601
Conduct Problems				
Implementation Fidelity	0.07	0.11	0.64	.523
Months in Treatment	-0.05	0.05	-0.92	.363
Hyperactivity-Inattention				
Implementation Fidelity	0.05	0.13	0.41	.686
Months in Treatment	-0.05	0.06	-0.80	.426
Peer Problems				
Implementation Fidelity	-0.02	0.09	-0.24	.811
Months in Treatment	0.05	0.04	1.24	.231
Prosocial Relations				
Implementation Fidelity	-0.07	0.12	-0.54	.594
Months in Treatment	0.07	0.06	1.18	.242
Impact				
Implementation Fidelity	0.06	0.15	0.38	.707
Months in Treatment	0.06	0.07	0.86	.393
Estimates of Covariance				
	Estimate	SE	Wald Z	p
Total Score				
Clinician	12.66	18.19	0.70	.486
Emotional Symptoms*	--	--	--	--
Conduct Problems				
Agency	0.88	1.07	0.82	.412
Hyperactivity-Inattention				
Clinician	0.00	0.04	0.06	.952
Peer Problems				
Clinician	2.17	0.96	3.31	.001
Prosocial Relations				
Agency	0.26	0.52	0.49	.625
Impact				
Agency	1.35	0.95	1.41	.158

*2-level models failed to converge, ANOVA conducted.

Table 19

SDQ Parent-Report Deltas: Estimated Fixed Effects and Covariance Parameters

Fixed Effects				
	β	<u>SE</u> (β)	t	p
Total Score				
Implementation Fidelity	-0.39	0.42	-0.93	.356
Months in Treatment	0.01	0.22	0.05	.961
Emotional Symptoms				
Implementation Fidelity	-0.00	0.16	-0.01	.993
Months in Treatment	0.10	0.09	1.08	.286
Conduct Problems				
Implementation Fidelity	0.10	0.15	0.63	.534
Months in Treatment	-0.05	0.08	-0.62	.539
Hyperactivity-Inattention				
Implementation Fidelity	-0.06	0.16	-0.35	.731
Months in Treatment	-0.09	0.09	-0.99	.325
Peer Problems				
Implementation Fidelity	-0.23	0.14	-1.68	.099
Months in Treatment	0.06	0.07	0.87	.388
Prosocial Relations				
Implementation Fidelity	-0.25	0.13	-1.92	.061
Months in Treatment	-0.07	0.07	-0.98	.332
Impact				
Implementation Fidelity	0.06	0.15	0.38	.707
Months in Treatment	0.06	0.07	0.86	.393
Estimates of Covariance				
	<u>Estimate</u>	<u>SE</u>	<u>Wald Z</u>	p
Total Score				
Clinician	21.78	15.55	1.40	.161
Emotional Symptoms				
Agency	0.02	0.03	0.58	.560
Conduct Problems				
Agency	0.60	1.11	0.54	.587
Hyperactivity-Inattention				
Agency	0.02	0.03	0.58	.560
Peer Problems				
Clinician	0.49	2.82	0.17	.862
Prosocial Relations				
Agency	0.00	0.02	0.18	.856
Impact				
Agency	1.35	0.95	1.41	.158

Table 20

*Client, Clinician, and Agency Characteristics Associated with Trauma-Focused
Cognitive-Behavioral Therapy (TF-CBT) Implementation Fidelity*

<u>Characteristic</u>	<u>Fidelity Components</u>				<u>Fidelity</u>
	<u>Content</u>	<u>Frequency</u>	<u>Duration</u>	<u>Coverage</u>	<u>Total</u>
<u>Clients</u>					
Age	-0.02	-0.03	-0.00	-0.06**	-0.04
Gender	0.00	0.00	0.04	0.02	0.02
Race	0.01	-0.02	0.06	-0.02	0.03
Ethnicity	-0.01	0.03	0.12**	-0.02	0.04
Primary Residence	0.04	0.08	0.06	-0.06**	0.06
Total # Trauma	0.01	0.03	-0.02	0.06**	0.00
<u>Types</u>					
Interpersonal Trauma	-0.08*	-0.01	-0.03	0.07**	-0.07 [†]
<u>Clinicians</u>					
Gender	-0.02	0.10*	-0.04	0.01	0.00
Degree	0.01	-0.03	-0.01	0.01	0.01
License	0.09*	-0.03	0.12**	0.02	0.09*
Supervisor Status	0.09*	0.03	0.05	0.08**	0.08*
<u>Agencies</u>					
Type	0.03	0.10	0.10*	0.03	0.08 [†]
Number of Sites	0.01	0.01	-0.02	0.01	0.03

** $p < .01$, * $p < .05$, [†] $p < .10$

Comprehensive Discussion

Although significant progress has been made in the development of trauma-focused EBTs for children and adolescents (Kolko et al., 2012), the transfer of these interventions from academic to community practice settings with robust fidelity and meaningful youth and family outcomes remains challenging (Forgatch & DeGarmo, 2011). Consequently, the continued investigation of large-scale implementation efforts for trauma-focused EBTs is imperative to ensure their effectiveness in settings where trauma-exposed populations are primarily served. This investigation is one of the first to examine the large-scale implementation of TF-CBT in community practice settings and adds to the TF-CBT implementation literature by studying nine years of client and implementation outcome data from 2,980 youth, 596 clinicians, and 75 agencies.

The first study, “The Large-Scale Implementation of Trauma-Focused Cognitive-Behavioral Therapy in Community Practice Settings: An Examination of Client Outcomes,” assessed the trauma history characteristics of clients participating in TF-CBT as well as client outcomes in the context of a statewide TF-CBT implementation initiative. Within community practice settings, findings suggested that almost half of clients participating in TF-CBT reported exposure to traumatic loss, bereavement, or separation; impaired caregiving; or witnessing domestic violence. Moreover, 85.5% of clients reported experiencing at least one, 66.7% at least three, and 40.8% at least five types of traumatic events. These trauma exposure rates are higher than rates discovered in NCTSN clinic-based service center samples across school-based, juvenile justice, mental health, intensive mental health, child welfare, and health centers (Briggs et al., 2013),

indicating that this sample was more highly traumatized than other NCTSN service center samples.

Given the heterogeneity of trauma exposure, and the fact that traumatic events rarely are single events, findings underscore the importance of an early, comprehensive trauma history assessment to assist clinicians in selecting an appropriate trauma-focused EBT or adjunctive treatment module (Briggs et al., 2013). Such assessments must determine the types, number, developmental periods, and density of trauma exposures (Briggs et al., 2013, p. 106), particularly given the burgeoning evidence that youth exposed to multiple, compared to single, traumas are at risk for subsequent trauma exposure and “cumulative impairment (e.g., psychiatric and addictive disorders; chronic medical illness; legal, vocational, and family problems)” (Cook et al., 2005, p. 390). If a clinician is trained in TF-CBT, for example, a comprehensive trauma history will assist in choosing appropriate adjunctive treatment modules, including those for child traumatic grief (Cohen et al., 2006; Cohen et al., 2017) or ongoing exposure to traumatic events (Cohen et al., 2011).

A second key finding from the first study indicated that TF-CBT participation was associated with statistically significant reductions in PTSD symptomatology from baseline to follow-up, as well as self- and parent-reported reductions in emotional symptoms, conduct problems, hyperactivity-inattention, peer problems, and the overall impact of these difficulties from baseline to follow-up. Youth also demonstrated statistically significant self- and parent-reported improvement in prosocial relations from baseline to follow-up. Effect sizes for change in client outcomes from baseline to follow-up were in the large range for PTSD symptomatology; in the medium range for parent-

reported total emotional and behavioral symptoms, impact of difficulties, emotional symptoms; and self-reported total emotional and behavioral symptoms. Notably, these effect sizes are comparable to outcomes reported by a smaller TF-CBT implementation initiative within a single community mental health agency (see Webb et al., 2014).

More often than not, efficacious interventions fail to achieve the same magnitude of effect once they are transported from academic “ivory tower” settings to more heterogenous community practice settings. This decrease in effect size is often referred to as “voltage drop” (Chambers et al.; 2013; Kilbourne et al., 2007). Given the large-scale nature of this statewide implementation initiative, the maintenance of TF-CBT’s association with reduced PTSD symptomatology from small to large-scale implementation research is very promising. Chambers and colleagues (2013) suggest that interventions often remain maximally beneficial when they are subjected to ongoing tailoring, evaluations, and refinement for disparate populations and settings. While TF-CBT tailoring and refinement were not formally evaluated during the Ambit Network’s nine-year TF-CBT implementation initiative, it is proffered that TF-CBT trainers and implementation administrators may have harnessed their understanding of trauma-focused EBT delivery in the state of Minnesota leading to enhanced implementation and sustainability.

The second study, “The Large-Scale Implementation of Trauma-Focused Cognitive-Behavioral Therapy (TF-CBT) in Community Practice Settings: An Assessment of Implementation Fidelity,” examined whether clinician implementation fidelity to TF-CBT was associated with changes in youth posttraumatic stress, emotional problems, and behavioral problems (i.e., conduct problems, hyperactivity-inattention,

peer problems) from baseline to follow-up. Multilevel modeling indicated that implementation fidelity scores were not significantly related to change in posttraumatic stress symptomatology, emotional problems, or behavioral problems as reported by youth and their caregivers. Although relationships between TF-CBT implementation fidelity and youth outcomes were not significant, associations were generally in the anticipated direction—greater clinician-reported implementation fidelity was related to greater change in youth symptomatology. However, for caregiver-reported outcomes, this relationship was contrary to expectations: greater TF-CBT implementation fidelity was related to the worsening of total emotional and behavioral problems, peer problems, and hyperactivity-inattention.

The caregiver-report findings were initially surprising given previous research on the impact of implementation fidelity and client outcomes suggesting that interventions with better, as opposed to poorer, implementation fidelity demonstrate effect sizes two to three times larger (Durlak & DuPre, 2008). It may be that clinicians' rigid adherence to TF-CBT protocols reduced flexibility and adaptability to client characteristics (Breitenstein et al., 2010), although such an interpretation is discordant with self-report findings. The literature on child-caregiver symptom report agreement provides a second interpretation in that a number of studies have found that parents tend to underestimate their children's PTSD symptomatology, particularly externalizing behaviors associated with PTSD (Ceballo et al., 2001; Shemesh et al., 2005; Stover et al., 2010).

Consequently, in this study, it may be possible that clinicians with better implementation fidelity included parents in TF-CBT on a more frequent basis, thus providing better psychoeducation about the effects of trauma and symptoms of PTSD. Increased caregiver

participation in TF-CBT may have improved their understanding and recognition of primary and secondary effects of trauma and resulted in increased caregiver insight into their child's internalizing and externalizing symptoms. In turn, increased insight may be associated with an overestimation of PTSD symptomatology.

In order to better understand trauma-focused EBT implementation fidelity in the context of community practice settings, the second study also investigated the client-, clinician-, and agency-level factors associated with TF-CBT implementation fidelity. Clinician license and supervisory status were significantly associated with total implementation fidelity score. Specifically, clinicians with professional counseling licenses (e.g., Licensed Professional Counselors, Licensed Professional Clinical Counselors), and those who provided supervisory services, provided TF-CBT with greater overall levels of implementation fidelity.

To generate a more nuanced understanding of client, clinician, and agency factors and their relationship with TF-CBT implementation fidelity, specific components of implementation fidelity were also assessed. Specifically, client exposure to interpersonal trauma, clinician license, and clinician supervisory status were significantly related to TF-CBT content; clinician gender and license were significantly related to frequency of implementation; client ethnicity, clinician supervisory status, and agency type were significantly related to treatment duration; and client age, primary residence, number of total trauma types, and exposure to interpersonal trauma were related to TF-CBT coverage (i.e., treatment completed as planned).

Results from the second study provide the TF-CBT implementation literature with a foundation upon which to build. Indeed, in accordance with previous implementation

research, TF-CBT implementation fidelity and its components were affected by factors located at multiple levels of the implementation context (see Damschroder et al., 2009; Durlak & DuPre, 2008; Fixsen et al., 2005), and these factors interacted in idiosyncratic and unpredictable ways (Damschroder et al., 2009). For example, it is difficult to reconcile the findings from this study that clients' exposure to interpersonal trauma led clinicians to implement TF-CBT with somewhat less fidelity, but with significantly greater odds of completion. It may be the case that youth with histories of severe interpersonal trauma are more likely to be residing outside the home, for instance, in treatment foster care, residential treatment, or juvenile detention settings. Cohen and colleagues (2016) argue, for example, that providing trauma-focused EBTs to adjudicated youth in residential treatment facilities is advantageous for multiple reasons. For instance, adjudicated youth are available to participate in treatment in a long-term, ongoing, and structured manner, with fewer factors impacting attendance or duration of treatment (e.g., transportation, scheduling difficulties). Additionally, Cohen et al. posited that residential treatment facilities may feel safer than the external environment—whether at home or in the community—thus providing a therapeutic environment more conducive to, and supportive of, successful trauma narration and processing (Cohen et al., 2016, p. 157).

Two key considerations for future TF-CBT implementation research must be noted. First, future research should investigate the degree or magnitude of implementation fidelity needed to maintain TF-CBT effects in community practice settings (Breitenstein et al., 2013). Such research should differentiate among the four components of implementation fidelity (i.e., content, coverage, duration, frequency) and identify acceptable levels of each component that ensure positive youth and family

outcomes (Breitenstein et al., 2013). Such research may assist in revealing the degree to which TF-CBT, or trauma-focused EBTs more broadly, can be tailored, or adapted, to meet the needs of target populations in heterogeneous routine practice settings, including outpatient, inpatient, residential treatment, and juvenile detention agencies (Breitenstein et al., 2013).

Second, implementation initiatives like the one offered by the University of Minnesota's Ambit Network are time- and resource-intensive. Indeed, results from an investigation of 17 statewide TF-CBT dissemination and implementation initiatives estimated that, on average, statewide implementation efforts cost approximately \$500,000 per year (Sigel, Benton et al., 2013). Moreover, the average number of clinicians trained over a period of three to five years was approximately 250 (Sigel, Benton et al., 2013). With evidence suggesting that 60 percent of children and adolescents are exposed to child maltreatment or other traumatic events before the age of 17 (Finkelhor, Turner, Ormrod, & Hamby, 2009; Finkelhor, Turner, Shattuck, & Hamby, 2015), such costly initiatives training such few clinicians will not reach the youth and families most in need of trauma-focused treatment. It is not necessarily the TF-CBT treatment components that are the primary concern but, rather, the limited accessibility and affordability of TF-CBT training and utilization that maintain the high burden of trauma exposure and posttraumatic stress across the United States (Kazdin & Rabbitt, 2013). Consequently, additional research on alternative models of TF-CBT, or other trauma-focused EBT, delivery is needed.

In summary, this investigation examined the nine-year statewide implementation of TF-CBT in a variety of community practice settings. Although TF-CBT has been

found to be effective across diverse clinical settings, the study of its transportability has been impeded by the complexity and multifactorial nature of the implementation process. Presenting data from 2,980 clients, 596 clinicians, and 75 agencies, this study is one of the first and the largest to examine TF-CBT client and implementation practice outcomes. Findings suggest that the widescale implementation of TF-CBT using the learning collaborative model is a promising approach to increasing the availability of trauma-focused EBTs in routine practice settings, while also reducing the primary and secondary effects of trauma in the lives of children and their families.

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 Services Research*, 38(1), 4-23.
- Aarons, G. A., James, S., Monn, A. R., Raghavan, R., Wells, R. S., & Leslie, L. K. (2010). Behavior problems and placement change in a national child welfare sample: A prospective study. *Journal of the American Academy of Child & Adolescent Psychiatry*, 49(1), 70-80.
- Aarons, G. A., Sommerfeld, D. H., & Walrath-Greene, C. M. (2009). Evidence-based practice implementation: The impact of public versus private sector organization type of organizational support, provider attitudes, and adoption of evidence-based practice. *Implementation Science*, 4(1), 83.
- Allen, B., & Johnson, J. C. (2012). Utilization and implementation of Trauma-Focused Cognitive–Behavioral Therapy for the treatment of maltreated children. *Child Maltreatment*, 17(1), 80-85.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: American Psychiatric Association.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Association.
- American Psychological Association Task Force on Evidence-Based Practice for

- Children and Adolescents. (2008). *Disseminating evidence-based practice for children and adolescents: A systems approach to enhancing care*. Washington, DC: American Psychological Association.
- Barber J. P., Sharpless, B. A., Klostermann, S., & McCarthy K. S. (2007). Assessing intervention competence and its relation to therapy outcome: A selected review derived from the outcome literature. *Professional Psychology: Research and Practice, 38*(5), 493-500.
- Bauer, M. S., Damschroder, L., Hagedorn, H., Smith, J., & Kilbourne, A. M. (2015). An introduction to implementation science for the non-specialist. *BMC Psychology, 3*(1), 32.
- Beers, S. R., & De Bellis, M. D. (2002). Neuropsychological function in children with maltreatment-related posttraumatic stress disorder. *American Journal of Psychiatry, 159*(3), 483-486.
- Bigfoot, D. S., & Schmidt, S. R. (2010). Honoring children, mending the circle: Cultural adaptation of trauma-focused cognitive-behavioral therapy for American Indian and Alaska Native children. *Journal of Clinical Psychology, 66*(8), 847-856.
- Bond, L., Glover, S., Godfrey, C., Butler, H., & Patton, G. C. (2001). Building capacity for system-level change in schools: lessons from the Gatehouse Project. *Health Education & Behavior, 28*(3), 368-383.
- Breitenstein, S. M., Gross, D., Garvey, C. A., Hill, C., Fogg, L., & Resnick, B. (2010). Implementation fidelity in community-based interventions. *Research in Nursing & Health, 33*(2), 164-173.
- Briggs, E. C., Fairbank, J. A., Greeson, J. K., Layne, C. M., Steinberg, A. M., Amaya-

- Jackson, L. M., ... & Pynoos, R. S. (2013). Links between child and adolescent trauma exposure and service use histories in a national clinic-referred sample. *Psychological Trauma: Theory, Research, Practice, and Policy*, 5(2), 101.
- Bruns, E. J., Kerns, S. E., Pullmann, M. D., Hensley, S. W., Lutterman, T., & Hoagwood, K. E. (2015). Research, data, and evidence-based treatment use in state behavioral health systems, 2001–2012. *Psychiatric Services*, 67(5), 496-503.
- California Evidence-Based Clearinghouse for Child Welfare. (2008). *Trauma-focused cognitive behavioral therapy—detailed report*. Retrieved from <http://www.cebc4cw.org/program/trauma-focused-cognitive-behavioral-therapy/>.
- Carlson, V., Cicchetti, D., Barnett, D., & Braunwald, K. (1989). Disorganized/disoriented attachment relationships in maltreated infants. *Developmental psychology*, 25(4), 525.
- Carroll, K., Nich, C., & Rounsaville, B. (1998). Utility of therapist session checklists to monitor delivery of coping skills treatment for cocaine abusers. *Psychotherapy Research*, 8(3), 307-320.
- Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science*, 2(40), e1-e9.
- Cary, C. E., & McMillen, J. C. (2012). The data behind the dissemination: A systematic review of trauma-focused cognitive behavioral therapy for use with children and use. *Children and Youth Services Review*, 34, 748-757.
- CATS Consortium. (2007). Implementing CBT for traumatized children and adolescents

- after September 11: Lessons learned from the Child and Adolescent Trauma Treatments and Services (CATS) Project. *Journal of Clinical Child and Adolescent Psychology*, 36, 581-592.
- Ceballo, R., Dahl, T. A., Aretakis, M. T., & Ramirez, C. (2001). Inner-city children's exposure to community violence: How much do parents know? *Journal of Marriage and Family*, 63(4), 927-940.
- Chadwick Center for Children and Families. (2004). *Closing the quality chasm in child abuse treatment: Identifying and disseminating best practices*. San Diego, CA.
- Chambers, D. A., Glasgow, R. E., & Stange, K. C. (2013). The dynamic sustainability framework: addressing the paradox of sustainment amid ongoing change. *Implementation Science*, 8(1), 117.
- Chaudoir, S. R., Dugan, A. G., & Barr, C. H. (2013). Measuring factors affecting implementation of health innovations: a systematic review of structural, organizational, provider, patient, and innovation level measures. *Implementation Science*, 8(1), 22.
- Cicchetti, D., & Rogosch, F. A. (2001). The impact of child maltreatment and psychopathology on neuroendocrine functioning. *Development and Psychopathology*, 13(4), 783-804.
- Cicchetti, D., Rogosch, F. A., Lynch, M., & Holt, K. (1993). Resilience in maltreated children. *Development and Psychopathology*, 5(04), 629-647.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Cohen, J. A., Berliner, L., & Mannarino, A. (2010). Trauma focused CBT for children

with co-occurring trauma and behavior problems. *Child Abuse & Neglect*, 34(4), 215-224.

Cohen, J. A., Deblinger, E., Mannarino, A. P., & Steer, R. A. (2004). A multisite randomized controlled study of sexually abused, multiply traumatized children with PTSD: Initial treatment outcome. *Journal of the American Academy of Child and Adolescent Psychiatry*, 43, 393-402.

Cohen, J. A., Mannarino, A. P., & Deblinger, E. (2006). *Treating trauma and traumatic grief in children and adolescents*. New York, NY: The Guilford Press.

Cohen, J. A., Mannarino, A. P., & Deblinger, E. (2017). *Treating trauma and traumatic grief in children and adolescents (2nd Ed.)*. New York, NY: The Guilford Press.

Cohen, J. A., Mannarino, A. P., & Iyengar, S. (2011). Community treatment of posttraumatic stress disorder for children exposed to intimate partner violence: a randomized controlled trial. *Archives of Pediatrics & Adolescent Medicine*, 165(1), 16-21.

Cohen, J. A., Mannarino, A. P., Jankowski, K., Rosenberg, S., Kodya, S., & Wolford, G. L. (2016). A randomized implementation study of trauma-focused cognitive behavioral therapy for adjudicated teens in residential treatment facilities. *Child Maltreatment*, 21(2), 156-167.

Cohen, J. A., Mannarino, A. P., Kliethermes, M., & Murray, L. A. (2012). Trauma-focused CBT for youth with complex trauma. *Child Abuse & Neglect*, 36(6), 528-541.

Cohen, J. A., Mannarino, A. P., & Murray, L. K. (2011). Trauma-focused CBT for youth who experience ongoing traumas. *Child Abuse & Neglect*, 35(8), 637-646.

- Cohen, J. A., Mannarino, A. P., & Staron, V. R. (2006). A pilot study of modified cognitive-behavioral therapy for childhood traumatic grief (CBT-CTG). *Journal of the American Academy of Child & Adolescent Psychiatry*, 45(12), 1465-1473.
- Cook, A., Spinazzola, J., Ford, J., Lanktree, C., Blaustein, M., Cloitre, M., ... & Mallah, K. (2017). Complex trauma in children and adolescents. *Psychiatric Annals*, 35(5), 390-398.
- Copeland, W. E., Keeler, G., Angold, A., & Costello, E. J. (2007). Traumatic events and posttraumatic stress in childhood. *Archives of General Psychiatry*, 64(5), 577-584.
- Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1(1), 16-29.
- Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implementation Science*, 4(1), 50.
- D'Andrea, W., Ford, J., Stolbach, B., Spinazzola, J., & van der Kolk, B. A. (2012). Understanding interpersonal trauma in children: why we need a developmentally appropriate trauma diagnosis. *American Journal of Orthopsychiatry*, 82(2), 187-200.
- Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clinical Psychology Review*, 18(1), 23-45.

- de Arellano, M. A. R., Lyman, D. R., Jobe-Shields, L., George, P., Dougherty, R. H., Daniels, A. S., ... & Delphin-Rittmon, M. E. (2014). Trauma-focused cognitive-behavioral therapy for children and adolescents: Assessing the evidence. *Psychiatric Services, 65*(5), 591-602.
- Deblinger, E., Cohen, J., Mannarino, A. P., Hanson, R., & Runyon, M. (2005). *TF-CBT practiced checklist—self-report*. Unpublished instrument.
- Deblinger, E., Cohen, J., Mannarino, A., Murray, L., & Epstein, C. (2007). *TF-CBT practice checklist*. Retrieved from <https://tfcbt.org/tf-cbt-brief-practice-checklist/>.
- Deblinger, E., Mannarino, A. P., Cohen, J. A., & Steer, R. A. (2006). A follow-up study of a multisite, randomized, controlled trial for children with sexual abuse-related PTSD symptoms. *Journal of the American Academy of Child & Adolescent Psychiatry, 45*(12), 1474-1484.
- Deblinger, E., McLeer, S. V., Atkins, M. S., Ralphe, D., & Foa, E. (1989). Post-traumatic stress in sexually abused, physically abused, and nonabused children. *Child Abuse & Neglect, 13*(3), 403-408.
- Deblinger, E., Pollio, E., & Dorsey, S. (2015). Applying trauma-focused cognitive behavioral therapy in group format. *Child Maltreatment, 17*, 1-15.
- Deblinger, E., Steer, R. A., & Lippmann, J. (1999). Two-year follow-up study of cognitive behavioral therapy for sexually abused children suffering post-traumatic stress symptoms. *Child Abuse & Neglect, 23*(12), 1371-1378.
- Derzon, J. H., Sale, E., Springer, J. F., & Brounstein, P. (2005). Estimating intervention effectiveness: Synthetic projection of field evaluation results. *Journal of Primary Prevention, 26*(4), 321-343.

- Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*(3-4), 327.
- Ebert, L., Amaya-Jackson, L., Markiewicz, J., & Burroughs, J. (2008). The NCCTS learning collaborative model for the adoption & implementation of evidence-based mental health treatment: NCCTS guidelines for conducting a learning collaborative. *Los Angeles, CA, and Durham, NC: National Center for Child Traumatic Stress and Duke University Evidence-Based Practice Implementation Center.*
- Ebert, L., Amaya-Jackson, L., Markiewicz, J. M., Kisiel, C., & Fairbank, J. A. (2012). Use of the Breakthrough Series Collaborative to support broad and sustained use of evidence-based trauma treatment for children in community practice settings. *Administration and Policy in Mental Health, 39*, 187-199.
- Ehlers, A., Grey, N., Wild, J., Stott, R., Liness, S., Deale, A., ... & Manley, J. (2013). Implementation of cognitive therapy for PTSD in routine clinical care: effectiveness and moderators of outcome in a consecutive sample. *Behaviour Research and Therapy, 51*(11), 742-752.
- Ehrenreich-May, J., Southam-Gerow, M. A., Hourigan, S. E., Wright, L. R., Pincus, D. B., & Weisz, J. R. (2011). Characteristics of anxious and depressed youth seen in two different clinical contexts. *Administration and Policy in Mental Health and Mental Health Services Research, 38*(5), 398-411.
- Elhai, J. D., Layne, C. M., Steinberg, A. M., Brymer, M. J., Briggs, E. C., Ostrowski, S.

- A., & Pynoos, R. S. (2013). Psychometric properties of the UCLA PTSD reaction index. Part II: investigating factor structure findings in a national clinic-referred youth sample. *Journal of Traumatic Stress, 26*(1), 10-18.
- Ellis, P. D. (2009). Effect size calculators. Retrieved from <https://www.polyu.edu.hk/mm/effectsizafaqs/calculator/calculator.html>.
- Fairbank, J. A., Putnam, F. W., & Harris, W. W. (2007). The prevalence and impact of child traumatic stress. In M. J. Friedman, T. M. Keane, & P. A. Resick (Eds.), *Handbook of PTSD: Science and Practice* (pp. 229-251). New York, NY: The Guilford Press.
- Finkelhor, D., Turner, H., Ormrod, R., & Hamby, S. L. (2009). Violence, abuse, and crime exposure in a national sample of children and youth. *Pediatrics, 124*(5), 1411-1423.
- Finkelhor, D., Turner, H. A., Shattuck, A., & Hamby, S. L. (2015). Prevalence of childhood exposure to violence, crime, and abuse: Results from the national survey of children's exposure to violence. *JAMA pediatrics, 169*(8), 746-754.
- Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Institute.
- Fixsen, D. L., Blase, K. A., Duda, M. A., Naoom, S. F., & Van Dyke, M. (2010). Implementation of evidence-based treatments for children and adolescents: Research findings and their implications for the future. In J. R. Weisz & A. E.

- Kazdin (Eds.), *Evidence-based psychotherapies for children and adolescents* (2nd Ed.) (pp. 435-450). New York, NY: Guilford Press.
- Forgatch, M. S., & DeGarmo, D. S. (2011). Sustaining fidelity following the nationwide PMTO™ implementation in Norway. *Prevention Science, 12*(3), 235-246.
- Fritz, R. M., Tempel, A. B., Sigel, B. A., Conners-Burrow, N. A., Worley, K. B., & Kramer, T. L. (2013). Improving the dissemination of evidence-based treatments: Facilitators and barriers to participating in case consultation. *Professional Psychology: Research and Practice, 44*(4), 225.
- Garland, A. F., Kruse, M., & Aarons, G. A. (2003). Clinicians and outcome measurement: What's the use? *The Journal of Behavioral Health Services & Research, 30*(4), 393-405.
- Glisson, C., & Hemmelgarn, A. (1998). The effects of organizational climate and interorganizational coordination on the quality and outcomes of children's service systems. *Child Abuse & Neglect, 22*(5), 401-421.
- Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire (SDQ). *Journal of the American Academy of Child and Adolescent Psychiatry, 40*, 1337-1345.
- Greeson, J. K., Briggs, E. C., Kisiel, C. L., Layne, C. M., Ake III, G. S., Ko, S. J., ... & Fairbank, J. A. (2011). Complex trauma and mental health in children and adolescents placed in foster care: Findings from the National Child Traumatic Stress Network. *Child Welfare, 90*(6), 91-108.
- Hemmelgarn, A. L., Glisson, C., & Dukes, D. (2001). Emergency room culture and the

- emotional support component of family-centered care. *Children's Health Care, 30(2)*, 93-110.
- Henggeler, S. W., Melton, G. B., Brondino, M. J., Scherer, D. G., & Hanley, J. H. (1997). Multisystemic therapy with violent and chronic juvenile offenders and their families: the role of treatment fidelity in successful dissemination. *Journal of Consulting and Clinical Psychology, 65(5)*, 821-833.
- Hernandez, J. T., Lodico, M., & DiClemente, R. J. (1993). The effects of child abuse and race on risk-taking in male adolescents. *Journal of the National Medical Association, 85(8)*, 593-597.
- Hoven, C. W., Duarte, C. S., Lucas, C. P., Wu, P., Mandell, D. J., Goodwin, R. D., ... & Musa, G. J. (2005). Psychopathology among New York City public school children 6 months after September 11. *Archives of General Psychiatry, 62(5)*, 545-551.
- IBM Corp. (2014). IBM SPSS Statistics for Windows, Version 23.0. Armonk, NY: IBM Corp.
- Ibrahim, J. G., & Molenberghs, G. (2009). Missing data methods in longitudinal studies: A review. *Test, 18(1)*, 1-43.
- Johnson, R. M., Kotch, J. B., Catellier, D. J., Winsor, J. R., Dufort, V., Hunter, W., & Amaya-Jackson, L. (2002). Adverse behavioral and emotional outcomes from child abuse and witnessed violence. *Child Maltreatment, 7(3)*, 179-186.
- Kaplow, J. B., & Widom, C. S. (2007). Age of onset of child maltreatment predicts long-term mental health outcomes. *Journal of Abnormal Psychology, 116(1)*, 176-187.
- Katakoa, S., Zhang, L., & Wells, K. (2002). Unmet need for mental health care among

- U.S. children. *American Journal of Psychiatry*, 159(9), 1548-1555.
- Kazdin, A. E. (2002). The state of child and adolescent psychotherapy research. *Child and Adolescent Mental Health*, 7(2), 53-59.
- Kazdin, A. E., & Rabbitt, S. M. (2013). Novel models for delivering mental health services and reducing the burdens of mental illness. *Clinical Psychological Science*, 1(2), 170-191.
- Kilbourne, A. M., Neumann, M. S., Pincus, H. A., Bauer, M. S., & Stall, R. (2007). Implementing evidence-based interventions in health care: application of the replicating effective programs framework. *Implementation Science*, 2(1), 42.
- Kim, J., & Cicchetti, D. (2010). Longitudinal pathways linking child maltreatment, emotion regulation, peer relations, and psychopathology. *Journal of Child Psychology and Psychiatry*, 51(6), 706-716.
- Klimes-Dougan, B., August, G. J., Lee, C. Y. S., Realmuto, G. M., Bloomquist, M. L., Horowitz, J. L., & Eisenberg, T. L. (2009). Practitioner and site characteristics that relate to fidelity of implementation: The Early Risers prevention program in a going-to-scale intervention trial. *Professional Psychology: Research and Practice*, 40(5), 467-475.
- Knaup, C., Koesters, M., Schoefer, D., Becker, T., & Puschner, B. (2009). Effect of feedback of treatment outcome in specialist mental healthcare: meta-analysis. *The British Journal of Psychiatry*, 195(1), 15-22.
- Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking*. New York, NY: Springer.
- Kolko, D. J., Baumann, B. L., Herschell, A. D., Hart, J. A., Holden, E. A., & Wisniewski,

- S. R. (2012). Implementation of AF-CBT by community practitioners serving child welfare and mental health: A randomized trial. *Child Maltreatment, 17*(1), 32-46.
- Kolko, D. J., Cohen, J. A., Mannarino, A. P., Baumann, B. L., & Knudsen, K. (2009). Community treatment of child sexual abuse: A survey of practitioners in the National Child Traumatic Stress Network. *Administration and Policy in Mental Health and Mental Health Services Research, 36*(1), 37-49.
- Kolko, D. J., Iselin, A. M. R., & Gully, K. J. (2011). Evaluation of the sustainability and clinical outcome of Alternatives for Families: A Cognitive-Behavioral Therapy (AF-CBT) in a child protection center. *Child Abuse & Neglect, 35*(2), 105-116.
- Kolko, D. J., & Swenson, C. C. (2002). *Assessing and treating physically abused children and their families: A cognitive behavioral approach*. Thousand Oaks, CA: Sage.
- Lang, J. M., Franks, R. P., Epstein, C., Stover, Ck., & Oliver, J. A. (2015). Statewide dissemination of an evidence-based practice using Breakthrough Series Collaboratives. *Children and Youth Services Review, 55*, 201-209.
- Lang, J. M., Randall, K. G., Delaney, M., & Vanderploeg, J. J. (2017). A model for sustaining evidence-based practices in a statewide system. *Families in Society: The Journal of Contemporary Social Services, 98*(1), 18-26.
- Lansford, J. E., Dodge, K. A., Pettit, G. S., Bates, J. E., Crozier, J., & Kaplow, J. (2002). A 12-year prospective study of the long-term effects of early child physical maltreatment on psychological, behavioral, and academic problems in adolescence. *Archives of Pediatrics & Adolescent Medicine, 156*(8), 824-830.

- Lyon, A. R., Stirman, S. W., Kerns, S. E., & Bruns, E. J. (2011). Developing the mental health workforce: Review and application of training approaches from multiple disciplines. *Administration and Policy in Mental Health and Mental Health Services Research, 38*(4), 238-253.
- Masten, A. S. (2001). Ordinary magic: Resilience processes in development. *American Psychologist, 56*(3), 227.
- McHugh, R. K., & Barlow, D. H. (2010). The dissemination and implementation of evidence-based psychological treatments: A review of current efforts. *American Psychologist, 65*(2), 73.
- McMullen, J., O'Callaghan, P., Shannon, C., Black, A., & Eakin, J. (2013). Group trauma-focused cognitive behavioral therapy with former child soldiers and other war-affected boys in the DR Congo: A randomized controlled trial. *Journal of child Psychology and Psychiatry, 54*, 1231-1241.
- Medical University of South Carolina. (2005). TF-CBTWeb: A web-based learning course for Trauma-Focused Cognitive-Behavioral Therapy. Retrieved from <https://tfcbt.musc.edu/>.
- Merikangas, K. R., He, J. P., Brody, D., Fisher, P. W., Bourdon, K., & Koretz, D. S. (2010). Prevalence and treatment of mental disorders among US children in the 2001–2004 NHANES. *Pediatrics, 125*(1), 75-81.
- Mihalic, S. F., Fagan, A., & Argamaso, S. (2008). Implementing the LifeSkills Training during drug prevention program: Factors related to implementation fidelity. *Implementation Science, 3*, 1-16.
- Minnesota Department of Human Services. (2008). Request for proposals for a qualified

grantee(s) to develop clinical capacity to provide Trauma-Focused Cognitive Behavioral Therapy. Minnesota Department of Human Services, Children's Mental Health Division.

Murray, L. K., Cohen, J. A., Ellis, B. H., & Mannarino, A. (2008). Cognitive behavioral therapy for symptoms of trauma and traumatic grief in refugee youth. *Child and Adolescent Psychiatric Clinics of North America*, 17(3), 585-604.

Murray, L. K., Dorsey, S., Skavenski, S., Kasoma, M., Imasiku, M., Bolton, P., Bass, J., & Cohen, J. A. (2013a). Identification, modification, and implementation of an evidence-based psychotherapy for children in a low-income country: The use of TF-CBT in Zambia. *International Journal of Mental Health Systems*, 7(1), 24.

Murray, L. K., Familiar, I., Skavenski, S., Jere, E., Cohen, J., Imasiku, M., Mayeya, J., Bass, J. K., & Bolton, P. (2013b). An evaluation of trauma focused cognitive behavioral therapy for children in Zambia. *Child Abuse and Neglect*, 37(12), 1175-1185.

National Child Traumatic Stress Network. (2014). General Trauma Information Form. Duke University and the University of California, Los Angeles

Nezlek, J. B. (2008). An introduction to multilevel modeling for social and personality psychology. *Social and Personality Psychology Compass*, 2(2), 842-860.

Nilsen, P. (2015). Making sense of implementation theories, models and frameworks. *Implementation Science*, 10(1), 53.

North Shore—Long Island Jewish Health System, Inc. (2006). North Shore Trauma History Checklist. Great Neck, NY: NSLIJHS.

- O'Callaghan, P., McMullen, J., Shannon, C., Rafferty, H., & Black, A. (2013). A randomized-controlled trial of trauma-focused cognitive behavioral therapy for sexually exploited, war-affected Congolese girls. *Journal of the American Academy of Child and Adolescent Psychiatry, 52*, 359-369.
- Pickreign Stronach, E., Toth, S. L., Rogosch, F., Oshri, A., Manly, J. T., & Cicchetti, D. (2011). Child maltreatment, attachment security, and internal representations of mother and mother-child relationships. *Child Maltreatment, 16*(2), 137-145.
- Pollak, S. D., Cicchetti, D., Hornung, K., & Reed, A. (2000). Recognizing emotion in faces: Developmental effects of child abuse and neglect. *Developmental Psychology, 36*(5), 679-688.
- Proctor, E. K., Landsverk, J., Aarons, G., Chambers, D., Glisson, C., & Mittman, B. (2009). Implementation research in mental health services: An emerging science with conceptual, methodological, and training challenges. *Administration and Policy in Mental Health and Mental Health Services Research, 36*(1), 24-34.
- Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., ... & Hensley, M. (2011). Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health and Mental Health Services Research, 38*(2), 65-76.
- Pynoos, R. S., Fairbank, J. A., Steinberg, A. M., Amaya-Jackson, L., Gerrity, E., Mount, M. L., & Maze, J. (2008). The National Child Traumatic Stress Network: Collaborating to improve the standard of care. *Professional Psychology: Research and Practice, 39*(4), 389-395.
- Pynoos, R., N. Rodriguez, A. Steinberg, M. Stuber, and C. Frederick. (1998). ULA PTSD

- Reaction Index for DSM-IV. Los Angeles, CA: UCLA Trauma Psychiatry Service.
- Pynoos, R., & Steinberg, M. (2013). UCLA PTSD Reaction Index for DSM-5. Los Angeles, CA: UCLA Trauma Psychiatry Service.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Thousand Oaks: Sage Publications.
- Rohrbach, L. A., Dent, C. W., Skara, S., Sun, P., & Sussman, S. (2007). Fidelity of implementation in Project Towards No Drug Abuse (TND): A comparison of classroom teachers and program specialists. *Prevention Science*, 8(2), 125.
- Rohrbach, L. A., Graham, J. W., & Hansen, W. B. (1993). Diffusion of a school-based substance abuse prevention program: Predictors of program implementation. *Prevention Medicine: An International Journal Devoted to Practice and Theory*, 22, 237-260.
- Ruzek, J. I., & Rosen, R. C. (2009). Disseminating evidence-based treatments for PTSD in organizational settings: A high priority focus area. *Behaviour Research and Therapy*, 47(11), 980-989.
- Salloum, A., Scheeringa, M. S., Cohen, J. A., & Storch, E. A. (2014). Development of stepped care trauma-focused cognitive-behavioral therapy for young children. *Cognitive and Behavioral Practice*, 21(1), 97-108.
- Saunders, B. E., Berliner, L., & Hanson, R. F. (2004). Child physical and sexual abuse: Guidelines for treatment (revised report, pp. 61–68). *Charleston, SC: National Crime Victims Research and Treatment Center*.
- Scheeringa, M. S., Weems, C. F., Cohen, J. A., Amaya-Jackson, L., & Guthrie, D.

- (2011). Trauma-focused cognitive-behavioral therapy for posttraumatic stress disorder in three-through six-year-old children: A randomized clinical trial. *Journal of Child Psychology and Psychiatry*, 52(8), 853-860.
- Schoenwald, S. K., Garland, A. F., Chapman, J. E., Frazier, S. L., Sheidow, A. J., & Southam-Gerow, M. A. (2011). Toward the effective and efficient measurement of implementation fidelity. *Administration and Policy in Mental Health and Mental Health Services Research*, 38(1), 32-43.
- Schoenwald, S. K., & Hoagwood, K. (2001). Effectiveness, transportability, and dissemination of interventions: What matters when? *Psychiatric Services*, 52(9), 1190-1197.
- Shemesh, E., Newcorn, J. H., Rockmore, L., Shneider, B. L., Emre, S., Gelb, B. D., ... & Yehuda, R. (2005). Comparison of parent and child reports of emotional trauma symptoms in pediatric outpatient settings. *Pediatrics*, 115(5), e582-e589.
- Shirk, S. R., Karver, M. S., & Brown, R. (2011). The alliance in child and adolescent psychotherapy. *Psychotherapy*, 48(1), 17.
- Shonk, S. M., & Cicchetti, D. (2001). Maltreatment, competency deficits, and risk for academic and behavioral maladjustment. *Developmental Psychology*, 37(1), 3-17.
- Sigel, B. A., Benton, A. H., Lynch, C. E., & Kramer, T. L. (2013). Characteristics of 17 statewide initiatives to disseminate Trauma-Focused Cognitive Behavioral Therapy (TF-CBT). *Psychological Trauma: Theory, Research, Practice, and Policy*, 5(4), 323-333.
- Sigel, J. A., & Williams, L. M. (2003). The relationship between child sexual abuse and

- female delinquency and crime: A prospective study. *Journal of Research in Crime and Delinquency*, 40, 71-94.
- Silverman, W. K., Ortiz, C. D., Viswesvaran, C., Burns, B. J., Kolko, D. J., Putnam, F. W., & Amaya-Jackson, L. (2008). Evidence-based psychosocial treatments for children and adolescents exposed to traumatic events. *Journal of Clinical Child & Adolescent Psychology*, 37(1), 156-183.
- Simpson, D. D. (2002). A conceptual framework for transferring research to practice. *Journal of Substance Abuse Treatment*, 22(4), 171-182.
- Singal, A. G., Higgins, P. D., & Waljee, A. K. (2014). A primer on effectiveness and efficacy trials. *Clinical and Translational Gastroenterology*, 5(1), e45.
- Southam-Gerow, M. A., Rodríguez, A., Chorpita, B. F., & Daleiden, E. L. (2012). Dissemination and implementation of evidence based treatments for youth: Challenges and recommendations. *Professional Psychology: Research and Practice*, 43(5), 527.
- Southam-Gerow, M. A., Weisz, J. R., & Kendall, P. C. (2003). Youth with anxiety disorders in research and service clinics: Examining client differences and similarities. *Journal of Clinical Child and Adolescent Psychology*, 32(3), 375-385.
- 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(2), 96-100.
- 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), 1-9.
- Stover, C. S., Hahn, H., Im, J. J., & Berkowitz, S. (2010). Agreement of parent and child reports of trauma exposure and symptoms in the early aftermath of a traumatic event. *Psychological Trauma: Theory, Research, Practice, and Policy, 2*(3), 159.
- Tarullo, A. R., & Gunnar, M. R. (2006). Child maltreatment and the developing HPA axis. *Hormones and Behavior, 50*(4), 632-639.
- treatment, and law* (pp. 353-374). New York, NY: Elsevier Science.
- Turner, H. A., Finkelhor, D., & Ormrod, R. (2006). The effect of lifetime victimization on the mental health of children and adolescents. *Social Science & Medicine, 62*(1), 13-27.
- U.S. Department of Health and Human Services, Administration for Children and Families, Children's Bureau. (2011). *Child maltreatment 2010*. Washington, DC.
- von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., Vandenbroucke, J. P., & Strobe Initiative. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *PLoS Medicine, 4*(10), e296.
- Waller, G. (2009). Evidence-based treatment and therapist drift. *Behaviour Research and Therapy, 47*(2), 119-127.
- Webb, C., Hayes, A. M., Grasso, D., Laurenceau, J. P., & Deblinger, E. (2014). Trauma-focused cognitive behavioral therapy for youth: Effectiveness in a community setting. *Psychological Trauma: Theory, Research, Practice, and Policy, 6*(5), 555-562.

- Weiner, D. A., Schneider, A., & Lyons, J. S. (2009). Evidence-based treatments for trauma among culturally diverse foster care youth: Treatment retention and outcomes. *Children and Youth Services Review, 31*(11), 1199-1205.
- Weisz, J. R., & Addis, M. E. (2006). The Research-Practice Tango and Other Choreographic Challenges: Using and Testing Evidence-Based Psychotherapies in Clinical Care Settings. In C. D. Goodheart, A. E., Kazdin, & R. J. Sternberg (Eds.), *Evidence-based psychotherapy: Where practice and research meet* (pp. 179-206). Washington, DC: American Psychological Association.
- Weisz, J. R., & Gray, J. S. (2008). Evidence-based psychotherapy for children and adolescents: Data from the present and a model for the future. *Child and Adolescent Mental Health, 13*(2), 54-65.
- West, B. T. (2009). Analyzing longitudinal data with the linear mixed models procedure in SPSS. *Evaluation & the Health Professions, 32*(3), 207-228.
- Widom, C. S., Czaja, S. J., Bentley, T., & Johnson, M. S. (2012). A prospective investigation of physical health outcomes in abused and neglected children: new findings from a 30-year follow-up. *American journal of public health, 102*(6), 1135-1144.
- Woods-Jaeger, B., Kava, C., Akiba, C., Lucid, L., Dorsey, S., & Kendall-Tackett, Kathleen. (2017). The Art and Skill of Delivering Culturally Responsive Trauma-Focused Cognitive Behavioral Therapy in Tanzania and Kenya. *Psychological Trauma: Theory, Research, Practice, and Policy, 9*(2), 230-238
- Woody, J. D., Anderson, D. K., & D'Souza, H. J. (2015). Dissemination of trauma-

focused cognitive-behavioral therapy with community practitioners: focus on self-efficacy. *Journal of Evidence-Based Social Work*, 12(3), 323-335.

Woody, J. D., Anderson, D. K., D'Souza, H. J., Baxter, B., & Schubauer, J. (2015).

Dissemination of trauma-focused cognitive-behavioral therapy: A follow-up study of practitioners' knowledge and implementation. *Journal of Evidence-Based Social Work*, 12(3), 289-301.

Yoo, J., Brooks, D., & Patti, R. (2007). Organizational constructs as predictors of effectiveness in child welfare interventions. *Child Welfare*, 86(1), 53-78.