


ORIGINAL ARTICLE

Measuring fidelity of empirically-supported treatment foster care: Preliminary psychometrics of the together facing the challenge—fidelity of implementation test (TFTC-FIT)

Dannia G. Southerland¹ | Elizabeth M.Z. Farmer²  | Maureen E. Murray¹ |
Leyla F. Stambaugh³ | Rachel D. Rosenberg²

¹Duke University School of Medicine,
Durham, NC, USA

²Virginia Commonwealth University, School of
Social Work, Richmond, Virginia, USA

³RTI International, Research Triangle Park, NC,
USA

Correspondence

E. M.Z. Farmer, Virginia Commonwealth
University, School of Social Work, 1000 Floyd
Ave., Richmond, VA 23284
Email: efarmer4@vcu.edu

Funding information

National Institute of Mental Health, Grant/
Award Number: MH057448

Abstract

This manuscript describes the development and psychometric testing of an instrument to assess implementation fidelity of Together Facing the Challenge (TFTC), an evidence-based approach to treatment foster care (TFC). The development of TFTC utilized a hybrid approach by bringing together components from evidence-based interventions with data on “what works” in usual care treatment foster care to create an intervention designed to provide enhanced treatment across a wide range of treatment foster care programs. The resulting training/consultation model has shown positive results on practice and youth outcomes in a randomized trial. The TFTC-FIT was developed during this randomized trial to be used as both a research measure and a quality improvement tool by provider agencies. Findings show that the measure is psychometrically sound and significantly distinguishes agencies trained in TFTC from those that are not. Discussion highlights the potential for such a measure, both in treatment foster care and in other home-/community-based interventions and settings for youth.

KEYWORDS

evidence-based practices, foster care, mental health, research methods, residential care

1 | INTRODUCTION

Efforts to improve the quality and effectiveness of child mental health services have included a focus on the development and dissemination of evidence-based practices (Hoagwood, Burns, & Weisz, 2002). However, there is a gap between what we know from research on effective interventions and widespread implementation of such interventions (Fixsen, Blase, Naoom, & Wallace, 2009; Hoagwood et al., 2001; Schoenwald et al., 2008). Although there are many challenges to translating research into practice, it is impossible to even examine this issue without sound measurement approaches that make it possible to know whether and to what extent evidence-based interventions are being implemented in everyday practice. Therefore, there is a need for reliable and valid measures of implementation fidelity (Bond, Drake, McHugo, Rapp, & Whitley, 2009; Fixsen, Naoom, Blase, Friedman, & Wallace, 2005; Schoenwald et al., 2011; Schoenwald & Garland, 2013).

Fidelity in mental health services research is generally defined as a measure of adherence to the standards and principles of a program

model (Bond, Evans, Salyers, Williams, & Kim, 2000), and a fidelity scale quantifies the degree to which a program adheres to the intended model (Drake et al., 2001). The use of fidelity measures has served a number of purposes in the attempt to improve treatment quality and outcomes in children's mental health, from explication of the “black-box” (i.e., what actually happens in treatment), and identification of critical ingredients of practice to use as a tool to guide program implementation (Mowbray, Holter, Teague, & Bybee, 2003). Fidelity measurement has also helped to clarify and quantify implementation of interventions and has provided evidence on the relationship between fidelity and treatment outcomes (e.g., Bruns, Suter, & Leverentz-Brady, 2008; Forgatch, Patterson, & DeGarmo, 2005; Schoenwald, Brown, & Henggeler, 2000).

Along with assessing the structural and programmatic aspects of a model, fidelity measures can also be an important adjunct to more traditional quality assurance strategies (Fixsen et al., 2009). As a quality assurance tool, fidelity scales can provide a reliable structure for assessing changes in practice and the evolution of practice towards

the evidence-based model. Ongoing fidelity monitoring is the foundation of sustainability of new practices (Aarons, Hurlburt, & Horwitz, 2010). For instance, measures of fidelity used in a program based feedback system can identify changes that take place—or do not—in the behaviour of the practitioners implementing a new program.

This paper introduces a measure of model fidelity developed in conjunction with a randomized trial of an enhanced approach to treatment foster care (TFC), Together Facing the Challenge (TFTC; Farmer, Burns, Wagner, Murray, & Southerland, 2010; Murray, Southerland, Farmer, & Ballentine, 2010). The measure, called Together Facing the Challenge-Fidelity of Implementation Tool (TFTC-FIT) is designed to address two critical purposes. First, it is designed to assess the extent to which model implementation is consistent with core strategies and techniques of TFTC (i.e., relationship building and teaching cooperation, setting expectations, use of effective parenting tools, implementing effective consequences, preparing youth for future transitions, and taking care of self). Second, the measure is designed to be used as a quality assurance tool, to be used in ongoing continuous quality improvement efforts to maintain the integrity of the model by agencies utilizing it. The challenge of bringing these two purposes into alignment is to develop a psychometrically sound measure of model fidelity that is rigorous enough for research and evaluation purposes and straight-forward enough to be used in everyday practice as part of continuous quality improvement efforts (Schoenwald et al., 2011).

2 | BACKGROUND

TFC is a community-based residential intervention for youth with mental health or behavioural problems (Chamberlain, 1994; Chamberlain & Mihalic, 1998; Chamberlain, Leve, & DeGarmo, 2007; Farmer et al., 2010). TFC provides intensive individualized treatment within the context of a family and community setting. Trained foster parents (treatment parents) work with foster youth in their homes to provide a structured, therapeutic environment. TFC provides the least restrictive treatment-based residential option for youth with serious mental health or behavioural problems, meaning that youth are provided opportunities for interactions and development within a family- and community-based setting, while still receiving the intensive treatment they need.

Treatment foster care may be known by a variety of other names across agencies and localities (e.g., “therapeutic foster care,” “specialized foster care,” and “intensive foster care”). What distinguishes TFC from traditional foster care is its view of treatment foster parents as front-line professionals who are part of a treatment team and who are responsible for implementing a treatment plan, rather than as just parent/caregiver substitutes. In line with this view of treatment parents as professionals, they receive more extensive pre-service and in-service training than traditional foster parents, and receive more supervision around how to effectively work with youth in their homes. TFC is most commonly viewed as an appropriate placement for youth with behavioural, emotional, and/or mental health problems whose needs cannot be adequately met in traditional foster care. It is often viewed as a preferred alternative to group residential placements for youth with serious problems who may be referred from a variety of sources

(e.g., mental health, juvenile justice, child welfare, and individual families). Although TFC can be used for a wide range of youth, the current study included youth who ranged in age from 2 to 21 (mean age of approximately 13), were racially diverse, were 45% female, and were in the “clinical” range on both internalizing and externalizing disorders (Farmer et al., 2010). TFC's appeal comes from its ability to combine intensive intervention approaches with a community- and family-based setting, thereby combining contemporary emphases on individualized treatment with placement in a less-restrictive setting. At this point in time, TFC is one of few community-based residential treatment options for which there is a substantial evidence base (Chamberlain, 2003; Farmer, Burns, Dubs, & Thompson, 2002; Fisher & Chamberlain, 2000; Farmer et al., 2010). Outcomes for TFC have shown decreased symptoms, increased functioning, post-discharge placements to lower levels of restrictiveness, and sustainability of improvements across time (e.g., Chamberlain et al., 2007; Eddy, Whaley, & Chamberlain, 2004; Farmer, Wagner, Burns, & Richards, 2003; Farmer et al., 2010).

A series of studies of one model of TFC, multidimensional treatment foster care (known as MTFC or, more recently, as TFC-Oregon), have provided the leading evidence base and have carefully described the evidence-based model (Chamberlain & Mihalic, 1998; Chamberlain, 1994; Chamberlain, 2002). MTFC/TFC-Oregon is being implemented by sites in the United States and internationally. Although this is tremendously promising as with most treatment approaches, the vast majority of agencies providing TFC are not providing this evidence-based version. While the definitive total number of TFC programs is unknown, estimates from various sources suggest that there are at least 3,000 TFC programs that provide usual care TFC across the United States—approximately 95% of these are *not* using MTFC (TFC-Oregon) as their model of care.

For over a decade, a set of studies focused on usual care TFC (Farmer et al., 2002; Farmer et al., 2010; Farmer, Murray, Southerland, Wagner, & Burns, 2009) found that TFC in usual care practice is often based on a home-grown treatment model and does not conform closely to either MTFC/TFC-Oregon or to national standards of care (Farmer et al., 2002; FFTA, 1995; FFTA, 2013). However, analyses also suggest that a variety of factors derived from MTFC and/or standards of care, when they occur in usual care TFC, are associated with improved outcomes for youth. These include increased training for treatment parents, improved supervision/consultation for treatment parents, and better relationships between the treatment parents and youth (Farmer et al., 2009). In the course of an observational study of TFC in usual care practice, interviews with directors of TFC agencies in a statewide sample showed that most of the directors wanted to improve quality and outcomes in their TFC programs (Farmer et al., 2002). However, directors were very clear that improved practice had to be viable within existing resources and needed to support underlying values and treatment approaches (for example, there was substantial resistance to requiring points and levels systems; most agencies did not have resources to hire their own therapists/clinicians, so they were reliant on available community-based providers; a substantial sub-set of youth remained in TFC long-term).

On the basis of these results on what “works” in usual care TFC and incorporating input from agency directors, (TFTC) was developed

(Farmer et al., 2009; Murray et al., 2010). Development of TFTC employed a hybrid model to improve practice, building on current practices in existing agencies, and infusing key elements from evidence-based interventions (via training and consultation) to overcome observed deficits in usual care practice. The intervention focuses on training TFC supervisors and treatment parents on key elements of behaviour management and relationship building, in an effort to increase effective strategies for improving outcomes with youth. Nearly all treatment parents in usual care TFC receive at least some training in behaviour management strategies (often during pre-service training). However, previous findings suggested that this training was not sufficient and that treatment parents had difficulty in effectively and consistently applying the skills they were introduced to during training (Farmer et al., 2009).

Hence, TFTC provided more intensive training in behaviour management approaches to more effectively manage foster youths' problem behaviour and encourage positive behaviour. TFTC is described in detail elsewhere (Murray et al., 2010; Murray, Culver, Farmer, Jackson, & Rixon, 2014). In brief, the training consisted of an initial two-day training with TFC supervisors to introduce them to the curriculum that would be used with the treatment parents and to help them become comfortable with the material so that they could co-facilitate training with treatment parents and work individually with their caseload of treatment parents on homework assignments between weekly group treatment parent training sessions. Treatment parent training consisted of 12 hr of group-based training, delivered once a week for 6 weeks. The treatment parent training focused on teaching approaches to addressing behaviour and integrating positive practices into the therapeutic process in six core areas: relationship building and teaching cooperation, setting expectations, use of effective parenting tools, implementing effective consequences, preparing kids for future transitions, and taking care of self. A randomized trial of TFTC showed that youth in the intervention group (i.e., in agencies participating in TFTC) showed significant improvement (compared to the youth in the control group) on the three focal domains—mental health symptoms, behaviours, and strengths (Farmer et al., 2010).

3 | CURRENT STUDY

Data collection for the larger randomized trial included interviews with treatment parents and youth at regular intervals for up to 2 years. These interviews collected a wide range of information about core processes and intervention elements in the home. However, they did not concisely assess whether the treatment parent was implementing the approaches and strategies covered in TFTC training. Therefore, the TFTC-FIT was developed during this study to provide an assessment of the extent to which TFC parents were implementing the components of TFTC with youth in their homes. Because the goal was to create a tool that could be used both for research purposes in the future and for assessing implementation quality by programs, initial efforts to test/use the instrument were piloted by the research team and then were completed by TFC supervisors who had been trained both in the TFTC model and on how to utilize the TFTC-FIT. Development and psychometric testing of the TFTC-FIT was

conducted by the same group of university-based researchers who conducted the larger randomized trial. This group included individuals with extended experience working within TFC programs, training staff, and conducting community-based studies on children's mental health services.

4 | METHODOLOGY

4.1 | Scale development

The initial step in developing a measure of implementation fidelity is identification of critical components of the model and translating them into objective and measurable indicators of adherence (Mowbray et al., 2003). As described above, the TFTC model was built on a combination of existing practice elements from a descriptive study of TFC in usual care and selected elements from evidence-based interventions. This process of deriving critical program elements led to the identification of the six core components of the TFTC model (See Figure 1).

Developing the fidelity measure for the enhanced TFC model (TFTC) involved operationalizing these core components into measurable indicators. The central focus of those efforts was to operationalize the six core components of the enhanced TFC model into scale items that captured the critical elements and range of each component domain. The resulting conceptual framework for the fidelity measure is shown in Figure 1. Focus groups with agency administrators, practitioners, and other stakeholders that had been involved in the intervention phase of the trial were employed in this stage of refining the fidelity criteria. We then piloted these initial scale items in two intervention agencies that had been implementing the TFTC model for a little over a year.

Although the TFTC-FIT was developed as part of a research study and could be used as a research instrument, it was designed primarily to be used as a fidelity monitoring and quality improvement instrument within agencies. Hence, ratings were based on direct observation during home visits by agency supervisors. Each item is rated on a 5-point Likert scale (0–4), from “not at all true for the treatment parent” (0) to “very much true for the treatment parent” (4). For quality improvement purposes, scores on individual items could identify areas where a particular treatment parent(s) is/are doing well and domains where improvement is needed. In addition, an overall scale score can be constructed based on the mean of the responses to the 16 questions. Using the conventions adopted by the National Implementing Evidence-Based Practices Project (McHugo et al., 2007), we characterized mean fidelity scores as high, moderate, and low. The decision making process to anchor the scale and individual item scores was done in consultation with the TFC agencies planning to use the final scale in their internal quality improvement process. On the basis of these conventions and consultation with agencies using the scale, individual item ratings of 3 or 4 were considered high fidelity, a score of 2 was considered moderate fidelity, and scores of 0 or 1 were considered low fidelity. For the composite (total mean) score, scores of 3.0 and above were considered high, 2–2.9 was moderate, and less than 2 was considered low fidelity. The final fidelity scale is included as an appendix.

Core Components of TFTC

	Building Relationships	Setting Expectations	Using Effective Parenting Tools	Implementing Effective Consequences	Preparing Youth for the Future	Taking Care of Self
Relevant Items on TFTC-FIT	Builds a therapeutic relationship	Establishes and utilizes daily check-in	Addresses thoughts, feelings, and behaviors	Utilizes problem solving techniques	Teaches relevant life skills	Incorporates family fun time
	Teaches cooperation	Tracks positive and negative behavior	Uses effective communications	Gives effective instructions		Takes care of self
		Develops and implements behavior contracts	Uses praise to encourage positive behavior	Implements consequences		
		Establishes and fine-tunes house rules	Interrupts the conflict cycle			

FIGURE 1 Link between core components and fidelity scale items

4.2 | Data collection for examination of the TFTC-FIT

Data for the initial piloting and psychometric testing of the measure were collected from a subset of agencies participating in the randomized trial. Testing involved a set of pilot studies (to test feasibility and refine the measure) and an evaluation of the scale's psychometric properties. The pilot was conducted in two administrations: first with four supervisors (who rated 28 treatment parents) and then with six supervisors (who rated 30 treatment parents). The final instrument was then utilized by a sample of 15 supervisors (who rated 110 treatment parents) to assess psychometric qualities.

All participants for each phase of the study were drawn from agencies participating in the larger randomized trial. For all phases, a member of the study team trained supervisors on how to complete the TFTC-FIT. Pilot studies were conducted in two participating intervention-arm agencies with a convenience sample of supervisors who volunteered for the activity (beyond their participation in the randomized trial training and data collection). The final psychometric testing involved two agencies that were geographically close to the study headquarters (one intervention, one control). All supervisors in both of these agencies participated in the final psychometric assessment by completing the measure for each family they supervised during the one-month focal period.

4.3 | Psychometric analysis

Psychometric properties were examined for all iterations of the fidelity scales.

For each iteration, Cronbach's alpha was used to determine internal consistency of the scale. Item analysis was used to investigate which, if any, items caused the Cronbach's alpha to increase if deleted (Tabachnik & Fidell, 2001). Factor analysis of the scale was performed to examine the measure's factor structure. More specifically, exploratory factor analyses were conducted on the second and the final iterations of the scale items for the TFTC-FIT. Factors were extracted

using principal axis factoring (Floyd & Widaman, 1995; Kline, 2005). Direct oblimin rotation was used because the items were expected to overlap. The Kaiser-Meyer-Olkin (KMO) value was computed to determine sampling adequacy for conducting a factor analysis. The KMO measures the degree of common variance among the variables in a factor analysis, on a scale between 0 and 1. KMO scores above .80 are considered strong (Weinberg & Abramowitz, 2008). Finally, examination of the scree plot was employed to determine the number of factors (Costello & Osborne, 2005).

5 | RESULTS

5.1 | Pilot scale

The first pilot test included the initial iteration of the measure with 17 items. The mean score for the first pilot scale was 2.9 with a standard deviation of less than 1, indicating that the respondents rated fidelity overall in the moderate range. The Cronbach's alpha was .89, indicating very good internal consistency. An item analysis of each item's correlation with the rest of the scale was also conducted. This analysis showed that three items did not perform as well as the others, suggesting that either (a) the wording of the items needed to be revised because they were not capturing what they are intended to capture, or (b) the variables measured by these items did not represent an included aspect of fidelity.

A focus group with agency supervisors who had been involved with the development and administration of the fidelity scale provided input on rewording these items and overall scale formatting. Based on feedback from these groups, one of the items was dropped (as it seemed redundant) and two items were reworded (to clarify their meaning). The scale, with the reworded items, was administered to another intervention site as the second phase of the pilot testing. The mean score for the second pilot of the scale was 3.3 (standard deviation = 0.44). This suggests respondents rated fidelity, on average

in the high range on the revised scale. The Cronbach's alpha was .91, indicating improvement in internal consistency. In addition, the performance of the reworded items was improved, as individual item analysis did not find that any single item reduced the Cronbach's Alpha in this second iteration of the fidelity scale.

5.2 | Final scale

Examination of the psychometrics from the first two pilots was incorporated into a third and final revision of the fidelity scale. Agency supervisors and members of the research team provided input on rewording the items and scale format, including developing specific operational definitions of each item and specific instructions on how the scale was to be administered. Psychometrics was examined for the final 16-item instrument, based on ratings from 15 supervisors on 110 treatment parents. The overall mean score for the revised scale was 3.3 (standard deviation = 0.58), which suggests the respondents rated fidelity, on average, in the high range. The scale's Cronbach's Alpha was .95, indicating excellent internal consistency. The improved Cronbach's Alpha suggests that the items on the scale appear to "hang together" measuring a single, overall construct (in this case, fidelity). We also conducted an item analysis to examine if revisions of the scale based on previous piloting, with some items that did not perform well being revised or deleted, had been sufficient. Results revealed that no deleted items increased Cronbach's Alpha beyond .95, an improvement over previous versions.

We then performed an exploratory factor analysis of the scale. The KMO value was .93, indicating that conducting a factor analysis on the scale was valid. All factor loadings were above .50, as shown in Table 1. Examination of the scree plot indicated that this was a one factor solution, and given that the eigenvalue for Factor 1 was high (9.03) corroborated the one factor solution. These findings confirm that the revised scale measured one overall latent construct (DeVellis, 2003).

TABLE 1 Factor loadings from exploratory factor analysis of third version of TFTC-FIT

Scale items	Factor loading
a. Building a therapeutic relationship	0.725
b. Teaches cooperation	0.811
c. Establishes and effectively utilizes a daily check-in	0.733
d. Tracks positive and negative behaviour	0.625
e. Develops and implements behaviour contracts	0.723
f. Establishes and fine-tunes house rules	0.610
g. Addresses thoughts, feelings, and behaviour	0.876
h. Uses effective communication	0.813
i. Uses praise to encourage positive behaviour	0.729
j. Interrupts the conflict cycle	0.835
k. Utilizes problem solving	0.812
l. Gives effective instructions	0.835
m. Implements consequences	0.847
n. Plans for the future	0.697
o. Includes family fun time	0.713
p. Takes care of self	0.538

Given these psychometric findings, we proceeded with examining the discriminate validity of the scale, by comparing between-group means for a TFTC-trained agency and a control (i.e., non-TFTC) agency to determine if these sites systematically differed on fidelity to the core principles.

The mean overall scale score for the comparison agency was 3.18, compared with 3.44 for the intervention agency. These group means were significantly different (Mann-Whitney $U = .931$, two-tailed $p = .02$). We also examined item level differences between groups. For each pair, the z score (similar to the t statistic) and asymptotic significance (two-tailed) is shown in the last two columns in Table 2. The intervention agency respondents had higher mean scores on all 16 included items than their control agency counterparts. For eight of these items, the differences were significant. The intervention agency treatment parents were rated significantly higher on the following: (a) Effectively utilizes daily check-in; (b) develops/implements behaviour contracts; (c) understand thoughts, feelings, and behaviour; (d) uses praise to encourage positive behaviour; (e) interrupts the conflict cycle; (f) problem solving; (g) gives effective instructions; and (h) family fun time.

6 | DISCUSSION AND IMPLICATIONS FOR PRACTICE AND RESEARCH

Based on these findings, the TFTC-FIT appears to be a psychometrically sound measure that provides a potentially important tool for assessing the level of implementation fidelity of TFTC. The measure showed acceptable internal consistency and ability to differentiate a particular model of care (TFTC) from other more generic versions of TFC. Given current findings, the TFTC-FIT could be used by researchers and/or agency personnel to assess implementation of TFTC in practice. Given ratings for non-TFTC TFC programs, it also appears to provide reasonable assessment of fidelity of key domains that are core to a broad range of TFC programs and agencies. Hence, current findings allow the measure to be used with confidence in its psychometric properties for assessing practice implementation and variation.

It should be noted that the differences between the intervention and comparison sites, while significant, are not huge. This is likely due to several factors. First, both agencies were providing TFC. Therefore, because the measure is assessing components that are central to the practice of TFC, it is encouraging that the comparison group showed substantial implementation of the core elements. It is also instructive to look at which items showed significant differences between the two groups and which did not. Overall, all items had a higher mean for the intervention group than the comparison group, but for eight of 16, this difference was statistically significant. All of the components that showed significant differences were particularly emphasized in TFTC (e.g., daily check-in; understands thoughts, feelings, behaviour; interrupts the conflict cycle; problem solving; gives effective instructions; and family fun time). However, some core constructs (e.g., builds a therapeutic relationship, effective communication, implements consequences, and takes care of self) did not differ between groups. These areas of non-significance may suggest domains

TABLE 2 Intervention and control means for TFTC-FIT

	Intervention agency means (s.d.)	Control agency means (s.d.)	t for equality of means
a. Building a therapeutic relationship	3.56 (.61)	3.45 (.64)	-.81
b. Teaches cooperation	3.44 (.66)	3.24 (.71)	-1.41
c. Effectively utilizes daily check-in	3.52 (.62)	2.91 (.94)	-3.41**
d. Tracks positive and negative behaviour	3.41 (.61)	3.32 (.84)	-.60
e. Develops/implements behaviour contracts	3.03 (.88)	2.43 (1.0)	-2.95**
f. Establishes and fine-tunes house rules	3.26 (.62)	3.07 (.96)	-1.10
g. Addresses thoughts, feelings, and behaviour	3.53 (.71)	3.08 (.81)	-2.79*
h. Uses effective communication	3.41 (.78)	3.28 (.81)	-.82
i. Uses praise to encourage positive behaviour	3.62 (.60)	3.38 (.65)	-1.79
j. Interrupts the conflict cycle	3.32 (.73)	3.01 (.79)	-1.95
k. Utilizes problem solving techniques	3.38 (.70)	3.03 (.77)	-2.30*
l. Gives effective instructions	3.50 (.66)	3.14 (.71)	-2.54**
m. Implements consequences	3.44 (.70)	3.18 (.74)	-1.85 ⁺
n. Plans for the future	3.59 (.56)	3.35 (.71)	-1.72 ⁺
o. Includes family fun time	3.71 (.60)	3.39 (.78)	-2.07*
p. Takes care of self	3.29 (.76)	3.20 (.65)	-.68

* $p < .1$.* $p < .05$.** $p < .01$

where training and supervision efforts need to be modified and supplemented in future implementations of TFTC. They may also suggest, however, that regardless of model, treatment parents do a reasonable job of many of the key TFC elements (e.g., building relationships, teaching cooperation, tracking behaviour, establishing house rules, and using praise to encourage positive behaviour). This is encouraging. It also suggests the need to understand “good enough” implementation (i.e., what is the threshold for adequate implementation of an intervention to have the desired effects on youth outcomes?)

The TFTC-FIT was developed to assess implementation of TFTC in TFC. However, ratings across both TFTC-implementing sites and more generic TFC-implementing sites suggest that it captures meaningful variation across all approaches to TFC. In addition, examination of its domains suggests that the measure may have more generalizable utility for the field. The components listed in the TFTC-FIT are, for the most part, core elements of many interventions designed to work with challenging youth within a family context. They include key aspects of parent management training, trauma-informed care, and cognitive-behavioural approaches to improving behaviour and outcomes. Additional work is needed to examine the cross-intervention/setting utility of the TFTC-FIT, both within interventions focused specifically on mental health and more generally among foster and other family-based and residential interventions.

7 | LIMITATIONS AND OTHER CONSIDERATIONS

Despite notable strengths of the study (e.g., relatively large sample size, inclusion of input from providers and other community-based stakeholders), this study did have limitations. One of those limitations was the use of non-independent observers scoring treatment parents

on the TFTC-FIT scale. This choice was predicated on feasibility considerations. Given the two purposes of this scale, as a measure of implementation fidelity and to develop an ongoing quality assurance feedback system, the flexibility of using on-site supervisors to assess fidelity is a critical component of the TFTC-FIT. As Schoenwald and colleagues (2011) make clear, efforts to balance the effectiveness and efficiency of measurement of implementation fidelity must accommodate the twin dictates of contextual fit and psychometric soundness.

The reported efforts are also a first step in testing and utilizing the TFTC-FIT in practice and/or research. The next step in further refining and testing this fidelity measure is to assess predictive validity and inter-rater reliability. Predictive validity will examine whether fidelity to the enhanced program and training/coaching model predict improved outcomes for youth. The assumption here is that the treatment parent management skills are a key mediator between the process of TFTC and the child outcomes (Eames et al., 2009). Inter-rater reliability will be critical to assess whether the measure can be completed similarly by supervisors, external observers (such as researchers/evaluators). This would clarify its practical and robust utility as both a quality measurement/improvement instrument and research instrument. Finally, if these further analyses yield positive results, it will be critical to work with a variety of agencies to determine the optimal way for incorporating the TFTC-FIT into a model of data-driven practice.

8 | CONCLUSION

The TFTC-FIT was developed with the goal of serving dual purposes. The first was to measure the fidelity of front-line practitioners, in this case treatment parents, in carrying out the desired interventions as

defined in the training. The second was to provide a mechanism for data driven feedback and evaluation of the treatment parents' ability to maintain the integrity of the practices over time. Further, the fidelity scale can provide a roadmap for coaching treatment parents in areas of weakness or departure from the evidence-based model.

Implementation of evidence-based practices in community-based "usual care" settings is especially difficult (Mendel, Meredith, Schoenbaum, Sherbourne, & Wells, 2008). Due to factors such as heterogeneity of staff background and skill level, resource and time commitment burdens necessary to implement an evidence-based intervention with fidelity, and other factors both specific and non-specific to a community, it is not hard to see why dissemination of evidence-based practice into most community-based settings has proven particularly problematic (Fixsen et al., 2005). Developing behaviour-based fidelity measures for such interventions and subjecting these measures to rigorous psychometric testing are important steps to ensure implementation that is both faithful to program specifications and consistent across providers. Continued work in this area will also be critical to identify the intervention components that are most critical for achieving positive child outcomes in TFTC and other community-based and home-based intervention approaches.

ACKNOWLEDGEMENT

This work was partially funded by a grant from the National Institute of Mental Health (MH057448).

ORCID

Elizabeth M.Z. Farmer  <http://orcid.org/0000-0001-9480-5819>

REFERENCES

- Aarons, G., Hurlburt, M., & Horwitz, S. (2010). 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.
- Bond, G. R., Drake, R. E., McHugo, G. J., Rapp, C. A., & Whitley, R. (2009). Strategies for improving fidelity in the National Evidence-Based Practices Project. *Research on Social Work Practice*, 19(5), 569–581.
- Bond, G. R., Evans, L., Salyers, M. P., Williams, J., & Kim, H.-W. (2000). Measurement of fidelity in psychiatric rehabilitation. *Mental Health Services Research*, 2(2), 75.
- Bruns, E., Suter, J., & Leverenz-Brady, K. (2008). Is it wraparound yet? Setting quality standards for implementation of the wraparound process. *The Journal of Behavioral Health Services and Research*, 35(3), 240–252.
- Chamberlain, P. (1994). *Family connections: A treatment foster care model for adolescents with delinquency*. Eugene, OR: Castalia Publishing Company.
- Chamberlain, P. (2002). Treatment foster care. In B. J. Burns, & K. Hoagwood (Eds.), *Community treatment for youth: Evidence-based interventions for severe emotional and behavioral disorders* (pp. 117–138). New York, NY: Oxford University Press.
- Chamberlain, P. (2003). An overview of the history and development of the multidimensional treatment foster care model and the supporting research. In *Treating chronic juvenile offenders: Advances made through the Oregon multidimensional treatment foster care model* (pp. 47–67). DC, US: American Psychological Association.
- Chamberlain, P., Leve, L. D., & DeGarmo, D. S. (2007). Multidimensional treatment foster care for girls in the juvenile justice system: 2-year follow-up of a randomized clinical trial. *Journal of Consulting and Clinical Psychology*, 75(1), 187–193.
- Chamberlain, P., & Mihalic, S. (1998). *Blueprints for violence prevention, book eight: Multidimensional treatment foster care*. Boulder, Colorado: Center for the Study and Prevention of Violence.
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10(7), 1–9.
- DeVellis, R. F. (2003). *Scale development: Theory and applications*. Thousand Oaks, CA: Sage Publications.
- Drake, R. E., Goldman, H. H., Leff, H. S., Lehman, A. F., Dixon, L., Mueser, K., & Torrey, W. C. (2001). Implementing evidence-based practices in routine mental health service settings. *Psychiatric Services*, 5(2), 179–182.
- Eames, C., Daley, D., Hutchings, J., Whitaker, C. J., Jones, K., Hughes, J. C., & Bywater, T. (2009). Treatment fidelity as a predictor of behaviour change in parents attending group-based parent training. *Child: Care, Health and Development*, 35(5), 603–612.
- Eddy, J. M., Whaley, R. B., & Chamberlain, P. (2004). The prevention of violent behavior by chronic and serious male juvenile offenders: A 2-year follow-up of a randomized clinical trial. *Journal of Emotional and Behavioral Disorders*, 12, 2–8.
- Farmer, E. M., Burns, B. J., Dubs, M. S., & Thompson, S. (2002). Assessing conformity to standards for treatment foster care. *Journal of Emotional and Behavioral Disorders*, 10(4), 213–222.
- Farmer, E. M. Z., Wagner, H. R., Burns, B. J., & Richards, J. T. (2003). Treatment foster care in a system of care: Sequences and correlates of residential placements. *Journal of Child and Family Studies*, 12, 11–25.
- Farmer, E. M. Z., Burns, B. J., Wagner, H. R., Murray, M., & Southerland, D. G. (2010). Enhancing "usual practice" treatment foster care: Findings from a randomized trial on improving youths' outcomes. *Psychiatric Services*, 61(6), 555–561.
- Farmer, E. M. Z., Murray, M., Southerland, D., Wagner, R., & Burns, B. J. (2009). *Together Facing the Challenge: Preliminary findings from a randomized clinical trial of therapeutic foster care*. Paper presented at the 21st annual research conference proceedings, a system of care for children's mental health: Expanding the research base., Tampa, FL.
- Fisher, P. A., & Chamberlain, P. (2000). Multidimensional Treatment Foster Care: A program for intensive parenting, family support, and skill building. *Journal of Emotional and Behavioral Disorders*, 8(3), 156–164.
- Fixsen, D. L., Blase, K. A., Naoom, S. F., & Wallace, F. (2009). Core implementation components. *Research on Social Work Practice*, 19(5), 531–540.
- Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa, FL: Louis de la Parte Florida Mental Health Institute, National Implementation Research Network (FMHI Publication #231).
- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7(3), 286–299.
- Forgatch, M. S., Patterson, G. R., & DeGarmo, D. S. (2005). Evaluating fidelity: Predictive validity for a measure of competent adherence to the Oregon model of parent management training. *Behavior Therapy*, 36(1), 3–13.
- Foster Family-Based Treatment Association (1995). *Program standards for treatment foster care*. New York: Foster Family-Based Treatment Association.
- Foster Family-Based Treatment Association (2013). *Program standards for treatment foster care (revised)*. New York: Foster Family-Based Treatment Association.
- Hoagwood, K., Burns, B. J., Kiser, L., Ringeisen, H., & Schoenwald, S. K. (2001). Evidence-based practice in child and adolescent mental health services. *Psychiatric Services*, 52(9), 1179–1189.
- Hoagwood, K., Burns, B. J., & Weisz, J. R. (2002). A profitable conjunction: From science to service in children's mental health. In B. J. Burns, & K. Hoagwood (Eds.), *Community treatment for youth: Evidence-based interventions for severe emotional and behavioral disorders* (pp. 327–338). New York, NY: Oxford University Press, USA.

- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: The Guilford Press.
- McHugo, G. J., Drake, R. E., Whitley, R., Bond, G. R., Campbell, K., Rapp, C. A., ... Finnerty, M. T. (2007). Fidelity outcomes in the national implementing evidence-based practices project. *Psychiatric Services*, 58(10), 1279–1284.
- Mendel, P., Meredith, L., Schoenbaum, M., Sherbourne, C., & Wells, K. (2008). Interventions in organizational and community context: A framework for building evidence on dissemination and implementation in health services research. *Administration and Policy in Mental Health and Mental Health Services Research*, 35(1), 21–37.
- Mowbray, C. T., Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity criteria: Development, measurement, and validation. *American Journal of Evaluation*, 24(3), 315–340.
- Murray, M., Culver, T., Farmer, E. M. Z., Jackson, L. A., & Rixon, B. (2014). From theory to practice: One agency's experience with implementing an evidence-based model. *Journal of Child and Family Studies*, 23(5), 844–853.
- Murray, M., Southerland, D., Farmer, E., & Ballentine, K. (2010). Enhancing and adapting treatment foster care: Lessons learned in trying to change practice. *Journal of Child and Family Studies*, 19(4), 393–403.
- Schoenwald, S. K., Brown, T. L., & Henggeler, S. W. (2000). Inside multisystemic therapy. *Journal of Emotional and Behavioral Disorders*, 8(2), 113–127.
- Schoenwald, S. K., Chapman, J. E., Kelleher, K., Hoagwood, K. E., Landsverk, J., Stevens, J., ... The Research Network on Youth Mental Health. (2008). A survey of the infrastructure for children's mental health services: Implications for the implementation of empirically supported treatments (ESTs). *Administration and Policy in Mental Health and Mental Health Services Research*, 35(1), 84–97.
- Schoenwald, S. K., & Garland, A. F. (2013). A review of treatment adherence measurement methods. *Psychological Assessment*, 25(1), 146.
- 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, 32–43.
- Tabachnik, B., & Fidell, L. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA: Allyn and Bacon.
- Weinberg, S. L., & Abramowitz, S. (2008). *Statistics using SPSS: An integrative approach* (second ed.). London: Cambridge University Press.

How to cite this article: Southerland DG, Farmer EMZ, Murray ME, Stambaugh LF, Rosenberg RD. Measuring fidelity of empirically-supported treatment foster care: Preliminary psychometrics of the together facing the challenge—fidelity of implementation test (TFTC-FIT). *Child & Family Social Work*. 2017;1–8. <https://doi.org/10.1111/cfs.12415>

APPENDIX

Together Facing the Challenge—Fidelity of Implementation Tool

This scale is intended to be administered by the supervisor at the end of their in-home sessions with families on their caseload. Based on direct observation, rate the caregiver's ability (using the scale provided below)

to effectively implement the following parenting skills and techniques presented during the training on the "Together facing the Challenge" curriculum.

0 not at all true for the parent	1 somewhat true for the parent	2 moderately true for the parent	3 true for the parent	4 very much true for the parent
----------------------------------	--------------------------------	----------------------------------	-----------------------	---------------------------------

- a. Builds a therapeutic relationship—encourages and supports child by providing important building blocks in their relationship (e.g., genuine interest, identifying common ground, positive attitude, patient and understanding, and consistent and follow-through)
- b. Teaches cooperation—is able to balance use of corrective discipline within the context of a supportive environment.
- c. Establishes and effectively utilizes a daily check-in—has a set time each day, approximately 5–10 min; encourages child to talk openly and provide feedback on how they are doing; starts and ends with positive encouragement and “sandwiches in” any problems, issues, or concerns.
- d. Tracks positive and negative behaviour—observes specific child behaviours to learn more about the frequency, duration, and intensity of the behaviours.
- e. Develops and implements behaviour contracts—understands the rationale behind behaviour contracts and steps needed to develop an effective contract, able to implement and follow through with the plan.
- f. Establishes and fine-tunes house rules—has a clear set of house rules; limited in number, posted in a central location, and is reviewed regularly. The house rules are followed by all members of the household and provide a structure for a safe environment.
- g. Addresses thoughts, feelings, and behaviour—demonstrates ability to assist child in recognizing, talking, about, and dealing with difficult thoughts and feelings that emerge; helps the child to understand how their thoughts and feelings can impact their behaviour.
- h. Uses effective communication—has developed effective methods for open and on-going communication in family; structure has been clearly identified and presented to all members of family.
- i. Uses praise to encourage positive behaviour—positively reinforces child by finding opportunities to “catch” their child being good.
- j. Interrupts the conflict cycle—is able to identify conflicts that take place and demonstrates ability to effectively intervene by de-escalating the situation.
- k. Utilizes problem solving techniques—demonstrates ability to use the problem solving model to address a specific problem by defining it clearly, generating multiple solutions, and selecting best solution based on outcomes.
- l. Gives effective instructions—demonstrates ability to use effective strategies (i.e., is specific, clear, respectful, and follows through). Instructions are given in a calm manner with the youth's full attention.
- m. Implements consequences—demonstrates knowledge of and ability to effectively implement various forms of consequences for problem behaviours.
- n. Teaches relevant life skills—demonstrates ability to transform daily living activities into learning opportunities to assist youth in the development of independent living skills.
- o. Incorporates family fun time—is able to describe various fun activities that the family has participated in recently. Uses these opportunities on a regularly scheduled basis to enhance the quality of family relationships.
- p. Takes care of self—is able to recognize the impact that stress has on their life, the “warning signs” that make them aware of it, and the specific strategies they use to manage their stress level while taking time for self on a regularly scheduled basis.