Randomized Trial of PMTO in Foster Care: 12-Month Child Well-being, Parenting, and Caregiver Functioning Outcomes

Authors

Becci Akin, PhD, MSW[a]

Kyle Lang, PhD[b]

Yueqi Yan[c]

Thomas P. McDonald, PhD

[a] – University of Kansas School of Social Welfare, Lawrence KS, USA

[b] – Tilburg University, Tilburg, NL

[c] – Children and Families Futures, Lake Forest, CA, USA

Corresponding Author:

Becci A. Akin, PhD, MSW	Twente Hall, 1545 Lilac Lane
Assistant Professor	Lawrence, KS 66045
University of Kansas	Phone: 785-864-2647
School of Social Welfare	Email: beccia@ku.edu

Acknowledgement

This manuscript was part of the Kansas Intensive Permanency Project (KIPP), which was funded by the Children's Bureau, Administration on Children, Youth and Families, Administration for Children and Families, U.S. Department of Health and Human Services, under grant number 90-CT-0152. The article's contents are solely the responsibility of the authors and do not necessarily represent the official views of the Children's Bureau. The authors also wish to thank their community collaborators in this study: the Kansas Department for Children and Families, KVC Kansas, and Saint Francis Community Services, Inc.

Conflict of Interest

Declaration of interest: none.

Highlights

- All child social-emotional well-being outcomes showed linear improvement
- Results for parenting practices were inconclusive due to measurement issues
- 2 of 4 caregiver outcomes improved – mental health and readiness for reunification

Abstract

Evidence-supported parenting interventions (ESPIs) increasingly are used in child welfare to improve child well-being and parenting. However, little evidence exists on the effectiveness of ESPIs with biological families of children in foster care with serious behavioral health problems. To address this gap in the literature, we examined the outcomes of in-home Parent Management Training Oregon model (PMTO). PMTO was evaluated in a randomized trial in which 918 children were assigned to PMTO or services as usual with assessment at baseline, 6 months, and 12 months. Outcome domains included child social-emotional well-being, parenting, and caregiver functioning. Our results show PMTO demonstrated linear improvements in children's social-emotional functioning, problem behaviors, and social skills. Although results for parenting were inconclusive, two of four caregiver functioning outcomes (parent mental health and readiness for reunification) were significantly improved. Overall, these findings suggest PMTO and similar ESPIs may hold promise for promoting better outcomes for biological families of children in foster care with behavioral health problems.

Children's behavioral health has gained attention as a relevant factor for improving outcomes among children in foster care. A number of studies have documented that children's behavioral health problems are associated with a significantly lower likelihood of permanency (e.g., Authors, 2011, 2012; Connell, Katz, Saunders, & Tebes, 2006; Park & Ryan, 2009). In fact, prior studies have indicated that children with serious emotional and behavioral problems were nearly four times more likely to experience long-term foster care than their peers without these problems (Authors, 2012). Researchers have also revealed growing evidence for treatments to address behavioral health problems among children in foster care (Landsverk, Burns, Stambaugh, & Reutz, 2009; Pecora, Jensen, Romanelli, Jackson, & Ortiz, 2009; Romanelli et al., 2009). Likewise, policy makers and administrators have advanced the goal of child well-being, especially social-emotional well-being. In 2012, the Children's Bureau issued a memorandum on promoting children's social-emotional well-being, marking a shift toward emphasizing well-being in concert with safety and permanency (Children's Bureau, 2012).

One strategy to address children's behavioral health problems in child welfare is evidence-supported parenting interventions (ESPIs) (Barth, 2009; Barth et al., 2005; Biglan, 2014; Horwitz, Chamberlain, Landsverk, & Mullican, 2010). ESPIs have been promoted not only because of research that shows promise for improving child behavior problems, but also because of demonstrated effects that address other risk factors related to poor child welfare outcomes. For instance, studies have shown long-term cascading effects to include parental mental illness, parental substance abuse, and poverty (Barth, 2009; Patterson, Forgatch, & DeGarmo, 2010).

While ESPIs, and rigorous evaluations of them, have expanded to child welfare settings, most studies have not included biological parents of children in foster care (Linares, Montalto, Li, & Oza, 2006). To date, randomized studies of ESPIs in child welfare settings have included:

(1) foster parents (Dozier, Peloso, Lewis, Laurenceau, & Levine, 2008; Leve et al., 2012; Mersky, Topitzes, Grant-Savela, Brondino, & McNeil, 2016; Price et al., 2008); (2) biological parents who were selected with various criteria but may or may not have had their children in foster care (e.g., confirmed physical abuse, enrolled in head start and history of child protection reports) (Bernard et al., 2012; Chaffin, Funderburk, Bard, Valle, & Gurwitch, 2011; Chaffin et al., 2004; Hurlburt, Nguyen, Reid, Webster-Stratton, & Zhang, 2013; Lind, Bernard, Ross, & Dozier, 2014); or (3) biological parents whose children already reunified with them and were no longer in foster care (DeGarmo, Reid, Fetrow, Fisher, & Antoine, 2013; Oxford, Marcenko, Fleming, Lohr, & Spieker, 2016). When biological parents were included in randomized studies, most have been group-based ESPIs (Webster-Stratton & Reid, 2010), which may have different advantages and disadvantages than individual, home-based ESPIs, especially among families impacted by poverty and experiencing transportation-related barriers. Finally, this line of research has largely omitted studies of ESPIs that included adolescents, despite the fact that behavioral health problems are most prevalent among older children and teens in foster care (Burns et al., 2004; Leslie, Hurlburt, Landsverk, Barth, & Slymen, 2004).

To address these gaps in the evidence base, the present randomized trial examined a behavioral, ESPI that was tailored for children in foster care with serious emotional disturbance. Specifically, we investigated whether usual foster care services *plus* individual, in-home Parent Management Training, Oregon (PMTO) model would result in improvements in child well-being, parenting, and caregiver functioning among children and youth identified as having a serious emotional disturbance and between the ages of 3 and 16 years. The study's 6-month outcomes were previously reported, finding improvements in three measures of child social-emotional well-being, including caseworkers' assessments of social-emotional functioning and

parents' reports of problem behaviors and social skills (Authors, 2016) and four measures of caregiver functioning, including caseworkers' assessments of parents' mental health, substance use, use of social supports, and readiness for reunification (Authors, 2017). Regarding improvements in parenting practices, results were mixed. That is, an experimental design indicated no effects on parenting at post-test (Authors, 2017), and a quasi-experimental study using latent growth curve modeling indicated significant positive gains in parenting at post-test (Authors, 2017). The current study aims to investigate whether PMTO contributed to positive outcomes for child social-emotional well-being, parenting practices, and caregiver functioning at follow-up (i.e.,12-months post-baseline).

PMTO is a behavioral parent training program developed over the past four decades by the Oregon Social Learning Center (OSLC) and disseminated by its affiliate, Implementation Sciences International, Incorporated (ISII). PMTO represents one of a family of interventions the OSLC created and researched to address child and adolescent antisocial behavior (M. S. Forgatch & Patterson, 2010). These programs are based on social interaction learning theory (SIL), which asserts that children's problematic behaviors can be mediated with parenting because parents are the agents of change for affecting improvements in their children's behaviors (Patterson, 1982). PMTO's content focuses on teaching parents strategies to increase positive parenting practices and decrease coercive parenting (Forgatch & Gewirtz, 2017, p. 86). In the present study, PMTO was delivered in-home to individual families for up to 6 months.

This study represents a project funded under the U.S. Children's Bureau's Permanency Innovation Initiative (PII). This initiative was distinguished by the PII Approach (PII-TTAP & PII-ET, 2013), which provided a systematic framework for implementation and evaluation. Grantees worked with technical assistance providers to select, implement, and test a local service

strategy. Using a results-oriented accountability framework (Testa & Poertner, 2010), grantees moved progressively through stages of implementation and evaluation, and were required to satisfy certain criteria before passing the "tollgate" and proceeding to the next stage (PII-TAP & PII-ET, 2013). The present study describes intermediate outcomes of the summative evaluation conducted by local evaluators in one of the six federally-funded projects.

The intermediate outcomes of the summative evaluation were defined prospectively by the project's logic model (Authors, 2013) following a comprehensive exploration stage (Authors, 2012, 2014). Based on a multi-method, multi-informant approach, the results of the exploration stage identified key barriers to permanency as observed at the child, family, and system levels. The intervention, thus, was selected and tailored to address these key barriers (Authors, 2014). The project's theory of change describes the expected outcomes:

> Parents of children with SED face multiple problems that are complex in nature and not alleviated easily by current child welfare practice or within current child welfare timeframes. To bring about change of sufficient magnitude, resources must be dedicated to improve ineffective parenting practices and to connect parents with community resources and social supports, such as mental health and substance abuse treatment. When parenting and community connections are strengthened, a more adequate and pro-social environment for children is created. Moreover, when the family's interpersonal and social environment is bolstered, child functioning increases and behavior problems decrease. These changes combine to create readiness for family reunification, which leads to more timely and stable reunifications (Authors, 2014, p. x).

In brief, the logic model articulated improvements in short and intermediate outcomes related to children's functioning and behaviors, parenting, and caregiver functioning (mental health, substance use, social supports, and readiness for reunification). The expectation for changes in these outcomes were based on an extensive literature review on parenting interventions (Authors, 2014), including longitudinal randomized studies of PMTO that demonstrated similar outcomes (Patterson et al., 2010).

In this summative evaluation of the intermediate outcomes, we consider the extent to which the intervention affected the outcomes' development over the course of three measurement occasions. Rather than simply evaluating the treatment effect at one occasion or examining unstructured mean differences between measurement occasions, we hypothesized stable, linear change in the outcomes. If an intervention is maximally effective, improvement may continue after the intervention concludes (i.e., at the 6-month post-test). With three measurement occasions, a linear trend is the most reasonable structural model to consider because more complex trends (e.g., a quadratic trend) would perfectly replicate the observed means and, consequently, cannot be meaningfully tested. We hypothesized that outcomes would demonstrate linear growth within the intervention group. Given behavioral changes may require long-term observation (i.e., more than 12 months), we expected the linear trajectories to represent "healthier" patterns of development in the intervention group than in the comparison group. These "healthier" relative trajectories may manifest as larger increases (or smaller decreases) in desirable outcomes for the intervention group. Alternatively, they may indicate larger decreases (or smaller increases) in undesirable outcomes for the intervention group relative to the comparison group. Our models (described below) also provided estimates of the treatment effect at the final measurement occasion.

Method

Design

The study design was a two-group, longitudinal post-randomized consent trial. Post-randomized consent designs (aka, Zelen designs) are defined by randomization occurring first and participants consenting to the study after randomization and with knowledge of their group assignment (Zelen, 1990). One reason Zelen designs are used is to reduce potential bias in

randomized trials that may emerge when knowledge of the treatment influences study recruitment (e.g., people seek to participate because they are enthusiastic about the treatment) and outcomes (e.g., treatment group's expectations may inflate benefits and the control group's disappointment and demoralization may influence results) (Hinman et al., 2014). A second reason for Zelen design is to strengthen external validity, especially in situations where a treatment is applied to an entire population (e.g., families of children in foster care with serious emotional and behavioral problems). The idea is that Zelen designs may provide a more representative sample because an entire sample of eligible participants are included in the study, not just individuals who consent to participate in a randomized study. The design may offer more accurate estimates of total impact that can be expected from an intervention being introduced on a population because it helps account for non-compliance with service plans and no-show rates in the overall estimate of a treatment's benefits (Adamson, Cockayne, Puffer, & Torgerson, 2006).

Study enrollment occurred over a two-year period (September 2012–September 2014) in Kansas. The study was preregistered with the National Institute of Health at clinicaltrials.gov (NCT02152618). Human subjects' approval was provided by the University of Kansas Institutional Review Board. The PMTO developer, ISII, was involved in implementation of PMTO but did not participate in design, conduct, analysis, or reporting of this study.

Participants

The sample comprised children from all regions of the state who were between the ages of 3 and 16, entering or re-entering foster care, and identified as having emotional and/or behavioral problems within six months of this removal episode. As usual practice, caseworkers, who had been trained in using the Child and Adolescent Functioning Assessment Scale (CAFAS) and the

Preschool and Early Childhood Functional Scale (PECFAS) (Hodges & Wong, 1996; Hodges, Xue, & Wotring, 2004), administered the scales with all children entering foster care in accordance with child age. Children were identified as eligible when one of the following criteria were met: (1) for children 3-5 years old, a PECFAS score of 50 or higher, or a score of 20 on one subscale; (2) for children 6-16 years old, a CAFAS score of 60 or higher, or a score of 30 on one subscale; or, (3) had been identified by a Community Mental Health Center as having an SED; (4) had an Individual Education Plan for an emotional or behavioral disorder; (5) had a diagnosed mental disorder, and symptoms of that disorder were contributing to placement instability; (6) had a diagnosed mental disorder, a history of outpatient or inpatient mental health treatment, and was currently prescribed psychotropic medications; or, (7) had been admitted for inpatient psychiatric care within the last year. Additionally, family-level criteria were assessed by supervisors and included that at the time of study: 1) the child's case plan goal must be reunification; 2) parent must reside in the service area; 3) parent may not be incarcerated for longer than three months; and 4) parent cannot have a court-order of "no contact" with the child. Each case consisted of the identified child and identified parent, which included biological parents, stepparents, adoptive parents, or other adults serving as primary caregiver and whose child had been removed into state custody. The identified parent represented the caregiver with whom the child was to reunify at study enrollment. Among the 918 children allocated to the study, 102 had more than one parent seeking to reunify with them (e.g., divorced parents). For the present study, we selected the parent identified as the primary caregiver (i.e., parent who had spent the most time caring for the child on a daily basis).

Procedures

Data coordinators staffed within foster care agencies entered data on all eligible children into a secure web-based system, the Research Electronic Data Capture (REDCap) system (Harris et al., 2009). When two or more siblings met the eligibility criteria, one child was randomly selected from the sibling group. When children and parents were identified as eligible, a program supervisor contacted the parent to explain the study, ask for their consent to participate, and obtain written informed consent in-person. Children also assented to study participation in-person. Parents and youth received modest financial compensation for completing the assessments at each of three data collection waves (baseline (T1), 6-months post-test from baseline (T2), and 12-months follow-up from baseline (T3)). Data collection involved: (1) questionnaires administered to parents, (2) assessments completed by caseworkers, and (3) video-recorded family interaction tasks conducted with parent(s) and child. Data were collected in-person in families' homes or a private room in the foster care agency.

Randomization

Randomization procedures were semi-automated in REDCap. At the time of the study initiation, REDCap did not have a randomization module. Therefore, a university research assistant, who was in a separate unit from the evaluation team, developed a form and file-based procedure with multiple checks to ensure fidelity to a preset randomization assignment schedule. The preset schedule used a blocked randomization approach stratified by region and a 1:1 allocation ratio. The researcher managing the randomization process did not recruit families or interact with program staff and was the only person with access to the preset randomization schedule; thus, the allocation was concealed until after random assignment. No blinding was used after random assignment. Allocation of cases was conducted by university researchers who

were in a separate location from foster care staff. Participant recruitment and enrollment was led by trained supervisors of the foster care agencies and monitored by the project's principal investigator.

Intervention

PMTO. PMTO was delivered by the state's private contractors for foster care services across the state, comprising both urban and rural communities. The staffing model included six units, each with a full-time master's level supervisor per five full-time practitioners plus one half-time administrative assistant. Units were located strategically in six offices across the state. Three full-time data coordinators served the six sites. Frontline staff were master's level practitioners, most of whom were licensed social workers, about one quarter were licensed marriage and family therapists, and the other quarter were licensed counselors. PMTO training required practitioners to participate in 8 days (64 hours) of pre-service training followed by 10 additional days (80 hours) of training over approximately 8 months. Practitioners also participated in two full days (16 hours) of in-person coaching. In addition to this initial coaching, they received ongoing observation-based coaching twice per month in one of three formats: written feedback, live feedback via video-conference, and/or live feedback via group.

Fidelity to the PMTO model was monitored by trainers and coaches via videos of practitioners' work with families (see Authors, 2017 for additional description). All PMTO sessions were video recorded using a video camera on a tripod, uploaded to a secure portal, and could be selected for review by coaches and/or fidelity raters. Certification took an average of 22 months to accomplish and required practitioners, in addition to the training and coaching requirements, to demonstrate their adherence to the model with a minimum of five families, and submit four session videos that addressed specific PMTO content with families (i.e., skill

encouragement and effective discipline). To achieve certification, all four videos had to be scored at specified minimum thresholds. Given a low caseload requirement (i.e., 4-6 cases per practitioner), a six-month service duration for delivering PMTO, and that some of the certification content was delivered in later family sessions, certification was a relatively lengthy process. Once certified, practitioners were rated for fidelity annually. Certification was not required for practitioners to deliver PMTO; rather, they started serving families after pre-service training. This study's intervention group were served by 46 practitioners, 16 of whom were certified PMTO therapist by the end of the study period. The last available fidelity rating in the present study indicated fidelity scores met the minimum threshold for noncertified and certified practitioners (Authors, 2017).

PMTO was delivered in-home to individual families, focusing on parents as agents of change, and delivered for up to six months. Children were transported to the home by the foster care agency. The program did not require a specific number of sessions or weeks; rather, practitioners worked with families until they completed the PMTO curriculum. Families who were retained for six months but did not complete the curriculum were discharged from the parenting program at six months. Typically, practitioners met with families twice per week for approximately 60-90 minutes per session plus a mid-week check-in that lasted for 20-30 minutes. Weekly sessions followed a three-step process. First, practitioners met with parents without children present. Second, parents were expected to practice new skills (e.g., noticing positive behaviors, identifying one's emotions, using problem-solving steps, etc.), and practitioners followed-up with parents by phone or in-person to discuss weekly 'homework.' Third, practitioners conducted a family session with parents and children together, during which parents tried newly learned skills with the practitioner present and acting as a live coach.

The PMTO curriculum centered on teaching parents five core parenting practices: 1) positive involvement; 2) skill building; 3) supervision and monitoring; 4) problem-solving; and 5) appropriate discipline (M. S. Forgatch & Patterson, 2010). Practitioners were guided by a pre-defined and semi-structured session outline provided by ISII (For sample session outline see: M. S. Forgatch & Domenech Rodríguez, 2016). To address pervasive trauma in both children and parents, project leaders asked that the PMTO training incorporate trauma-specific content. In brief, the PMTO training emphasized trauma content, a focus on emotion regulation, and use of mindfulness techniques, which were integrated throughout the curriculum. Besides these modifications, PMTO did not undergo other major adaptations during the course of the study.

Other aspects of intervention fidelity were defined by the site implementers and not required by ISII. Based on findings from the exploration stage, other program components included low caseloads, early intervention, treatment completion, service duration and intensity, in-home service, and parent-child contacts (Authors, 2012, 2014, 2017).

Service as Usual. Participants randomized to the comparison group received services-as-usual from the foster care agencies. These services comprised case management delivered by staff with bachelor or master's degrees in a variety of fields and with varying levels of training. The state's foster care contracts required case managers to visit children and to arrange for at least one parent-child visit per week. Additional services offered to children and parents were not standardized in intensity or modality; rather, services were individualized to each family as determined by case managers, supervisors, court personnel, and other ancillary service providers.

Outcomes

Outcomes included three main categories of child outcomes, parenting, and caregiver functioning, which are reported for: baseline (T1), post-test at 6 months (T2), and follow-up at

12 months from baseline (T3). Below is a description of three variables for children's outcomes, one variable for effective parenting, and four variables for caregiver functioning.

Child outcomes. Child outcomes comprised social-emotional functioning, problem behaviors, and social skills. Below is a description of the measurement approach for each of these child outcomes.

Social-emotional functioning. Social-emotional functioning was measured using the CAFAS (ages 6-16) and the PECFAS (ages 3-5) (Hodges & Wong, 1996; Hodges et al., 2004), a caseworker-administered assessment (see Authors 2017 for additional information). The CAFAS provides an overall functioning score and eight subscales (School, Home, Community, Behavior Toward Others, Moods/Emotions, Thinking Problems, Self-Harm, and Substance Use). The PECFAS has seven subscales, omitting the substance use subscale. Scores were assigned via behaviorally oriented descriptions in increments of 10 where 0 = minimal functional impairment, 10 = mild functional impairment, 20 = moderate functional impairment, and 30 = severe functional impairment. Total scores represented sums of subscales and an overall level of functioning. Cronbach's alphas were estimated for each study group and each time point as follows: Intervention: T1 =. 64; T2 = .93; T3 = .92; Comparison: T1 = .61; T2 = .94; T3 = .93.

Child problem behaviors. The Social Skills Improvement System-Rating Scales (SSIS) (Gresham & Elliot, 1990) were used to assess child *problem behaviors* and *social skills* by administering parent versions, which were developed for ages 3 to 18 years. Data collection protocols required that the caregiver had had at least one visit with the child within the last 60 days. The SSIS provides two scores. First, it measures *problem behaviors* with a total score based on five subscales (33 items): externalizing, bullying, hyperactivity/inattention, internalizing, and Autism Spectrum. Second, the SSIS measures *social skills* (described below).

Parents were asked to report how often the child displayed the behavior on a 4-point scale (N = *never*, S = *seldom*, O = *often*, A = *almost always*). Higher problem behavior scores indicate more problem behaviors. Cronbach's alphas were estimated as follows: Intervention: T1 = .86, T2 = .86, T3 = .84, Comparison: T1 = .86, T2 = .81, T3 = .83.

Child social skills. As stated above, the SSIS also measured children's social skills. The scale provided a total score that comprises seven subscales (46 items): communication, *cooperation*, assertion, responsibility, empathy, engagement, and self-control. Like problem behaviors, parents were asked to report how often the child displayed the social skills on a 4-point scale (N = *never*, S = *seldom*, O = *often*, A = *almost always*). Higher social skills scores indicate stronger social skills. Cronbach's alphas were estimated as follows: Intervention: T1 = .91, T2 = .92, T3 = .93; Comparison: T1 = .90; T2 = .89; T3 = .92.

Effective parenting. Effective parenting was measured with the Family Interaction Task (FIT), which is an observation-based assessment that video-records the parent and index child working together on several tasks for approximately 30 minutes. The tasks are grouped into three developmentally-appropriate sets for preschool age children, school-age children, and adolescents. Videos were uploaded to a secure portal where they were observed and rated by coders. The coders were blind to the data collection wave and study condition, and were monitored by the study's principal investigator with regards to maintaining interrater reliability throughout the study. Reliability was checked on 15% of the sample and the percent agreement ranged from 66% to 98% with an average percent agreement of 89%. Coders rated behavioral items on their frequency according to these guidelines: never (0% of time), hardly ever (1-10% of time), sometimes (11-50% of time), often (51-75% of time), very often (76-90% of time), and always (91-100% of time). Some tasks sought specific practices or behaviors and these were

rated as: untrue (1), slightly true (2), fairly true (3), mostly true (4), and very true (5) (e.g., a problem solving task asked if several solutions were suggested and a plan was developed). While tasks and items within the age groupings of the FIT were specific to the child's developmental stage, all were rated and scored on five subscales (50 items) that correspond to the core parenting practices of PMTO: skill encouragement, positive involvement, problem-solving, communication/monitoring, and ineffective discipline. Subscales were reverse coded as needed (ineffective discipline) and averaged to provide an overall measure of effective parenting. Cronbach alphas were estimated for each study group and each time point as follows: Intervention: T1= .59; T2 = .70; T3 = .58; Comparison: T1 = .66; T2 = .68; T3 = .60.

Caregiver functioning. Four subscales (16 items) of the North Carolina Family Assessment Scale (NCFAS) were completed by case managers to represent caregiver functioning: parent mental health, parent substance use, parent use of social supports, and readiness for reunification (Reed-Ashcraft, Kirk, & Fraser, 2001). Scores were recorded with a six-point scale that ranged from "clear strength" (+2) to "serious problem" (-3) with anchoring definitions provided for three of the points (clear strength (+2), baseline/adequate (0), and serious problem (-3)).

Demographic and Case Characteristics

Child and parent demographic characteristics and placement information were included for baseline and attrition comparisons. These variables were collected from the state's administrative data which are reported semi-annually to the federally-required Adoption and Foster Care Reporting System (AFCARS) (see Table 1).

Analytic Strategy

The effect of the intervention was assessed using intention-to-treat (ITT) analyses. Where data were missing, multiple imputation was applied using the three waves of data. Missing data

at Time 2 and 3 were multiply imputed with 100 imputations using the principal component auxiliary variable technique (Howard, Rhemtulla, & Little, 2015) as implemented in the PcAux software (Lang, Little, & Team, 2018). This approach efficiently includes all possible predictor variables into the imputation models, thereby recovering the unobserved data structure as well as possible and correcting the nonresponse bias that simpler missing data treatments (e.g., listwise deletion, last observation carried forward) would fail to address (Lang & Little, 2018). Furthermore, this approach does not require (or benefit from) differentiating between attrition and item-level missing data during imputation.

In what we denote below as the Tier-1 analysis, we directly tested our hypotheses by modeling time as a linear trend (T1 = -2, T2 = -1, T3 = 0) and interacting time and intervention condition to test for moderation by intervention group (i.e., to model unique trends for the intervention and comparison groups). If the intervention condition did not significantly moderate the linear trend (i.e., if the estimated linear rate of change in the outcome was not significantly different between the intervention and comparison groups), we conducted a Tier-2 post-hoc analysis based on a more flexible model that dummy-coded measurement occasion with T1 set as the reference group. These models allowed us to estimate group-specific changes in the outcomes between baseline (T1) and the two follow-up occasions (T2 and T3). In all models, intervention was dummy coded (Intervention Group = 1, Comparison Group = 0). To fully describe group-specific trends, we conducted a simple slopes analysis. The results from the additional models used for the simple slopes analysis are provided in the supplemental file.

Results

Participant Flow and Attrition

Figure 1 illustrates the participant flow. Over the two-year recruitment period, 6,657 children were assessed for eligibility. After applying eligibility criteria, 1,652 children were identified as eligible and randomly assigned to the intervention (n = 855) or comparison (n = 797) groups. Among randomized cases, 394 intervention and 340 comparison cases were not approached due to timing and location of service openings. In total, 918 children were allocated to the study (intervention = 461 and comparison = 457). Consent rates were 78.1% and 66.5% for the intervention and comparison groups, respectively ($\chi^2 = 14.14$, $p < .001$). Further multivariate analyses of factors that predicted consent found no differences between the study groups (Authors, 2017). Attrition was specific to the measure and differed between the study groups (Authors, 2017), suggesting a missing at random (MAR) mechanism (i.e., the propensity to drop out depends on measured characteristics of the families; Enders, 2010). Given that data are MAR, simple missing data approaches, such as listwise deletion or last observation carried forward, would bias estimates of the intervention effect (Little & Rubin, 2014). Thus, we employed modern multiple imputation to avoid biased inferences that occur with MAR nonresponse (Rubin, 1987).

Intervention Fidelity

Multiple aspects of intervention fidelity were measured including adherence to the PMTO model, low caseloads, early intervention, treatment completion, service duration and intensity, in-home service, and parent-child contacts. All measures showed that adequate intervention fidelity was obtained (Authors, 2017). Additionally, reasons for intervention cases' non-completion were analyzed by Authors (2017).

Randomization and Comparability

To check randomization, approached ($N = 918$) and non-approached ($N = 734$) children were compared on 15 baseline variables. Approached and non-approached children were similar on all but two variables, race and age; however, the effect sizes were small to negligible (see Authors, 2017). Table 1 presents data for the intervention and comparison groups on 16 baseline variables, showing that they were similar on all variables.

Intervention Effects

Error! Reference source not found. shows the means and standard deviations for the study outcome variables at all three time points and for each study group. Additionally, Cohen's d indicates each group's effect size of the mean differences between T2 and T1, T3 and T2, and T3 and T1.

Child outcomes. Error! Reference source not found. presents the results from models for children's outcomes. The Tier-1 analysis showed that linear trends in social-emotional functioning, problem behaviors, and social skills were all significantly moderated by the intervention. For these three outcomes, our hypothesis was supported because the intervention group demonstrated "healthier" patterns of development than the comparison group. These conditional linear trends are plotted in Figures 2-4 in the supplemental file.

Social-emotional functioning. Time 3 (T3) CAFAS scores were significantly lower in the intervention group than in the comparison group ($\beta = -24.05$, $t = -5.12$, $p = 0.001$). In line with our hypothesis, the intervention group demonstrated stable CAFAS levels whereas the comparison group showed a significant worsening (i.e., increase) in CAFAS levels. The comparison group's linear trend in CAFAS scores was positive and significant ($\beta = 7.47$, $t =$

3.47, $p < 0.001$) whereas the intervention group showed a nonsignificant trend (β = -0.41, t = -0.21, p = 0.836).

Problem behaviors. The intervention group had significantly lower T3 levels of problem behaviors than the comparison group (β = -3.79, t = -3.86, p = 0.001). The comparison group's linear trend in problem behaviors was negative and significant (β = -2.10, t = -2.94, p = 0.003). The intervention group's linear trend was also significantly negative but had larger magnitude than the comparison group's trend (β = -4.47, t = -7.96, p <0.001). The significantly larger improvement (i.e., decrease) in problem behaviors for the intervention group, relative to the comparison group, supported our hypothesis.

Social skills. The intervention group had significantly higher levels of social skills (SS) at T3 (β = 5.81, t = 3.79, $p < 0.001$). The comparison group's linear trend in social skills was nonsignificant (β = -0.16, t = -0.17, p = 0.87), but the intervention group's linear trend was significant and positive (β = 2.96, t = 3.50, $p < 0.001$). Our hypothesis was, therefore, supported because the comparison group showed minimal change in social skills while the intervention group demonstrated a significant improvement (i.e., increase) in social skills.

Effective parenting. The measure for effective parenting, the FIT, was subject to large amounts of missing data that we were unable to adequately treat. Parameter estimates of FIT models, therefore, each had very high fraction of missing information (FMI) values (e.g., FMI = 0.85 for the interaction term in the moderated linear trends model), indicating that most of the sampling variability in the estimates was attributable to missing data. No effects in the Tier-1 moderated linear trend models or in the Tier-2 fully dummy-coded models achieved statistical significance (see Table 5 in supplemental file), but we cannot draw practical conclusions from

this lack of significance when considering the very high FMI values. Accordingly, we do not consider the FIT results in the following interpretations.

Caregiver functioning. The Tier-1 analysis showed that linear trends of caregiver functioning, as measured by NCFAS subscales, were not significantly moderated by the intervention (see Table 8 in supplemental file). That is, the best fitting linear trends did not significantly differ between intervention groups. Therefore, we estimated the Tier-2 descriptive models that used dummy codes to freely estimate changes from baseline within study group (Table 4). Conditional means from these models are plotted in Figures 5-8 in the supplemental file. Although we had no a priori hypotheses relating to the Tier-2 models, the general patterns of results suggested the that intervention had the desired effect on NCFAS outcomes at T2, but that this effect did not persist until T3—hence the lack of support for stable linear trajectories.

Mental health. Mental health ratings of the intervention group (M = -0.56) and comparison group (M = -0.46) were not significantly different at T1 (β = -0.10, t = -0.99, p = 0.33). Intervention group significantly moderated changes in mental health between T2 and T1 (β = 0.74, t = 4.80, p < 0.001) but not between T3 and T1 (β = 0.28, t = 1.77, p = 0.08). Intervention group's mental health ratings improved (i.e., increased) significantly from T1 to T2 (β = 0.21, t = 2.01, p = 0.04) and from T1 to T3 (β = 0.23, t = 2.16, p = 0.03) whereas the comparison group's mental health ratings worsened (i.e., decreased) significantly from T1 to T2 (β = -0.52, t = -4.79, p < 0.001) but were not different at T1 and T3 (β = -0.05, t = -0.43, p = 0.67).

Substance abuse. The substance abuse ratings of the intervention group (M = 0.003) and comparison group (M = -0.03) were not significantly different at baseline (β = 0.03, t = 0.27, p = 0.79). Intervention group significantly moderated differences in substance abuse ratings between T1 and T2 (β = 0.47, t = 2.81, p = 0.005) but not between T1 and T3 (β = 0.32, t = 1.80, p =

0.07). The intervention group's substance abuse ratings improved (i.e., increased) significantly between T1 and T2 (β = 0.27, t = 2.30, p = 0.02) but were not significantly different at T1 and T3 (β = -0.13, t = -1.03, p = 0.30). The comparison group's substance abuse ratings did not change significantly from T1 to T2 (β = -0.21, t = -1.71, p = 0.09) but significantly worsened (i.e., decreased) from T1 to T3 (β = -0.44, t = -3.52, p < 0.001).

Social support. Baseline support ratings of the intervention group (M = -0.32) and comparison group (M = -0.39) were not significantly different (β = 0.07, t = 0.83, p = 0.41). Intervention group significantly moderated changes in social support from T1 to T2 (β = 0.69, t = 4.99, p < 0.001) but not between T1 and T3 (β = 0.12, t = 0.81, p = 0.42). The intervention group's social support ratings improved (i.e., increased) significantly from T1 to T2 (β = 0.24, t = 2.47, p = 0.01) but were not significantly different at T1 and T3 (β = -0.04, t = -0.42, p = 0.68). The comparison's social support ratings worsened (i.e., decreased) significantly between T1 and T2 (β = -0.45, t = -4.59, p < 0.001) but did not differ at T1 and T3 (β = -0.16, t = -1.50, p = 0.14).

Readiness for reunification. Baseline readiness ratings of the intervention group (M = -0.92) and comparison group (M = -1.02) were not significantly different (β = 0.10, t = 0.92, p = 0.36). Intervention group significantly moderated changes in readiness between T1 and T2 (β = 0.42, t = 2.59, p = 0.01) but not between T1 and T3 (β = -0.12, t = -0.71, p = 0.48). The intervention group's readiness ratings improved (i.e., increased) significantly between T1 and T2 (β = 0.63, t = 5.65, p < 0.001) and between T1 and T3 (β = 0.44, t = 3.61, p < 0.001). The comparison group's readiness ratings did not significantly change between T1 and T2 (β = 0.21, t = 1.79, p = 0.07) but did improve (i.e., increase) significantly from T1 to T3 (β = 0.56, t = 4.58, p < 0.001).

Discussion

This study contributes to the question of whether an in-home, individual parenting intervention delivered to biological parents of children in foster care can improve the intermediate outcomes of children's social-emotional well-being, parenting, and caregiver functioning over time. Across most outcomes, this study's results suggest that children and families receiving the PMTO intervention fared better than those in the comparison group. In testing our hypothesis that the outcomes would demonstrate linear growth within the intervention group, it was supported for all child well-being outcomes. However, results on the parenting outcome were inconclusive due to poor measurement properties and highly influential missing data. Furthermore, while a within-group linear trajectories were not fully supported for caregiver functioning, several caregiver results suggest relatively healthier patterns of development in the intervention group. Below is a discussion of each outcome domain.

With regards to child social-emotional well-being outcomes, children in the PMTO intervention group showed healthier trajectories but three unique patterns emerged for social-emotional functioning, problem behaviors, and social skills, respectively. Findings on social-emotional functioning indicated that the intervention group remained stable from baseline to 12-month follow-up (i.e., no significant change). In contrast, the comparison group demonstrated small but significant worsening in social-emotional functioning. Regarding problem behaviors and social skills, the intervention group improved on both with the greatest effects on problem behaviors. The comparison group also improved on problem behaviors but the improvement was considerably smaller than the intervention group. Additionally, while the intervention group's growth in social skills may be considered small, the comparison group showed no changes in social skills.

The study's inconclusive findings on effective parenting warrant careful consideration. First, data collection procedures (i.e., video recording of parents and children together) were complicated due to participants' willingness to be video recorded and challenges with overcoming logistical barriers, particularly for children placed a long distance from their parents. These administration issues combined with the measure's scoring approach contributed towards a high amount of missing data, which we could not adequately address. Second, the estimation of Cronbach's alpha for the FIT showed low internal consistency, a possible red flag that the measure does not assess a single, unified construct of effective parenting. Third, other studies of similar interventions have found positive parenting results; however, most relied on parent self-reports and not on observation-based measures (Chamberlain et al., 2008; DeGarmo et al., 2013; Linares et al., 2006). While observational methods may be considered a gold standard in behavioral sciences, a meta-analysis of 33 studies cautioned their use with maltreating parents (Wilson, Rack, Shi, & Norris, 2008). In short, when interaction tasks were too short and undemanding, they did not detect aversive parenting behaviors. Thus, Wilson and colleagues recommended longer observation periods (minimum of 2 hours) and observation within natural settings, such as the family's home. Additionally, parenting differences were more difficult to detect in certain parent and child age groups, which represent the vast majority of the present study. Given these various issues, it seems possible that the FIT did not detect differences in parenting behaviors that may have existed. Although we have been able to implement latent growth curve analysis with data from a practitioner-administered measure and show improvements in parenting among the intervention group, these same data were not available for the comparison group (Authors, 2017). Thus, future research of parenting interventions in foster

care must address these measurement challenges and continue to examine parenting outcomes with rigorous designs.

The caregiver functioning domain included parent mental health, substance use, use of social supports and readiness for reunification. For all four outcomes, the results indicated that there was not a stable linear trend within the intervention group. Rather, the intervention group showed significant improvements at 6-months in all four outcomes but with a loss in these early gains at 12-months for two outcomes – substance use and social supports. Whereas the other two outcomes – mental health and readiness – also trended downward after initial improvements, the intervention group's 12-month scores were significantly better than baseline scores. In contrast, the comparison group's outcomes indicated either no significant change or significant worsening on all caregiver outcomes and every measurement occasion except for a small but significant change in readiness between baseline and 12-month follow-up. The comparison group's improvement in readiness at 12-months may not be surprising because parents in usual care services, like those in the intervention group, received case management services with a primary goal of reunification. These results suggest that workers considered usual care services as positively influencing readiness for reunification at 12-month follow-up.

In all, these findings contribute new evidence to the literature on parenting interventions in foster care settings. Based on a review of this literature, our study appears to represent the first to use a randomized design to examine an in-home parenting intervention with biological parents of children and adolescents in foster care. Among comparable existing studies, none included children over the age 12. To our knowledge, no other randomized studies within child welfare have investigated caregiver functioning outcomes. Our positive findings on children's social-emotional well-being are also important. Although prior studies have observed child behavior

change in the desired direction, they have not included a sample of children with emotional and behavioral problems specifically, nor have they demonstrated significant linear trends as in the present study (DeGarmo et al., 2013; Linares et al., 2006; Oxford et al., 2016). Additionally, while the follow-up caregiver functioning results are limited to positive effects on parent mental health and readiness for reunification, additional longitudinal study is needed to test for possible curvilinear relationships. Importantly, other PMTO research has described a "struggle work-through" in which parents make early improvements followed by a temporary setback; yet, in the long-term they realize greater gains (M. S. Forgatch & DeGarmo, 1999; Patterson & Chamberlain, 1992). Furthermore, a nine-year follow up study of PMTO with a non-clinical sample of single mothers found long-term advantages in parent functioning outcomes such as depression, employment, and police arrests (Patterson et al., 2010). More rigorous and long-range studies such as these are needed in child welfare to determine whether improvements emerge on a longer timeline and cascading effects occur in collateral outcome domains.

Limitations

Several study limitations should be noted. First, the research design requires consideration. Post-randomized consent designs are advantageous for reducing recruitment bias and strengthening external validity by extending results to an entire population of interest, not just individuals willing to participate in a randomized study (Adamson et al., 2006; Hinman et al., 2014). On the other hand, these designs may dilute intervention effects by including participants who did not indicate an interest to participate and, thereby, exacerbate attrition (Fan, 2015). While this study's attrition was sizable, we have reported it explicitly (Authors, 2017) and have addressed it with appropriate, modern statistical techniques. Given the complications of

conducting rigorous evaluations in child welfare settings, this design may optimize generalizability to target populations within foster care.

A second limitation pertains to a lack of blinding. Although the use of multiple outcome domains, multiple informants, and multiple measures may strengthen the study, a possible constraint includes a lack of blinding of researchers, parents, and case managers. Due to the logistics of foster care service delivery and court responsibilities, blinding was not possible.

Finally, interpretation of results should take into account that PMTO and services-as-usual were delivered by private, community-based foster care agencies under contract and monitoring by the state public child welfare agency. The effect of these organizational arrangements on generalizability is unknown. Similarly, two other potential confounders should be considered. First, this real-world evaluation setting did not allow control of the other services and supports received by children and parents. Despite the numerous other services that may have been provided, we have no reason to believe that they would have been allocated differentially to the study groups. Second, this study did not test the effect of parent-child contacts, which were likely greater for the intervention group due to concerted efforts to deliver PMTO. Future studies should examine whether an increase in parent-child contacts alone (without delivery of any curriculum) is associated with improved child social-emotional well-being, caregiver functioning, and parenting practices.

In conclusion, parenting interventions delivered in foster care settings appear to hold promise for improving outcomes, including when these interventions are delivered to biological parents of children in foster care. Further, these interventions may be applied to a wide age range of children and youth with behavioral health problems. Our 12-month follow up results suggest that an in-home parenting intervention delivered to biological families, even when the children are in

foster care and span a wide age range of 3 to 16 years, improved children's social-emotional well-being and some aspects of caregiver functioning. Future research should re-examine parenting outcomes with a robust measurement approach. Also necessary is longer-term study to determine whether parenting interventions lead to higher rates of family reunification. Given the importance of improving reunification outcomes, particularly for children and adolescents with behavioral health problems, researchers should strive to conduct rigorous and longitudinal studies on parenting interventions with biological families.

References

Adamson, J., Cockayne, S., Puffer, S., & Torgerson, D. J. (2006). Review of randomised trials using the post-randomised consent (Zelen's) design. *Contemporary Clinical Trials, 27*(4), 305-319. doi:10.1016/j.cct.2005.11.003

Barth, R. P. (2009). Preventing child abuse and neglect with parent training: Evidence and opportunities. *Future of Children, 19*(2), 95-118.

Barth, R. P., Chamberlain, P., Reid, J. B., Rolls, J., Hurlburt, M. S., Farmer, E. M. Z., . . . Kohl, P. L. (2005). Parent-training programs in child welfare services: Planning for more evidence-based approach to serving biological parents. *Research on Social Work Practice, 15*(5), 353-371. doi:10.1177/1049731505276321

Bernard, K., Dozier, M., Bick, J., Lewis-Morrarty, E., Lindhiem, O., & Carlson, E. (2012). Enhancing attachment organization among maltreated children: Results of a randomized clinical trial. *Child development, 83*(2), 623-636. doi:10.1111/j.1467-8624.2011.01712.x

Biglan, A. (2014). *A Comprehensive Framework for Nurturing the Well-being of Children and Adolescents*. Washington, DC: U.S. Deparment of Health and Human Services, Administration for Children Youth and Families, Children's Bureau

Burns, B. J., Phillips, S. D., Wagner, H. R., Barth, R. P., Kolko, D. J., Campbell, Y., & Landsverk, J. (2004). Mental health need and access to mental health services by youths involved with child welfare: A national survey. *Journal of the American Academy of Child & Adolescent Psychiatry, 43*, 960.

Chaffin, M., Funderburk, B., Bard, D., Valle, L. A., & Gurwitch, R. (2011). A combined motivation and parent-child interaction therapy package reduces child welfare recidivism in a randomized dismantling field trial. *Journal of Consulting and Clinical Psychology, 79*(1), 84-95. doi:10.7037/a0021227

Chaffin, M., Silovsky, J. F., Funderburk, B., Valle, L. A., Brestan, E. V., Balachova, T., . . . Bonner, B. (2004). Parent-Child Interaction Therapy with physically abusive parents:

Efficacy for reducing future abuse reports. *Journal of Consulting and Clinical Psychology, 72*(3), 500-510. doi:10.1037/0022-006X.72.3.500

Chamberlain, P., Price, J. M., Leve, L. D., Laurent, H., Landsverk, J. A., & Reid, J. B. (2008). Prevention of behavior problems for children in foster care: Outcomes and mediation effects. *Prevention Science, 9*(1), 17-27. doi:10.1007/s11121-007-0080-7

Children's Bureau. (2012). *Information Memorandum on Promoting Social and Emotional Well-Being for Children and Youth Receiving Child Welfare Services. ACYF-CB-IM-12-04*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Famlies. Retrieved from http://www.acf.hhs.gov/sites/default/files/cb/im1204.pdf

Connell, C. M., Katz, K. H., Saunders, L., & Tebes, J. K. (2006). Leaving foster care--the influence of child and case characteristics on foster care exit rates. *Children and Youth Services Review, 28*(7), 780-798. doi:10.1016/j.childyouth.2005.08.007

DeGarmo, D. S., Reid, J. B., Fetrow, B. A., Fisher, P. A., & Antoine, K. D. (2013). Preventing child behavior problems and substance use: The Pathways Home foster care reunification intervention. *Journal of Child & Adolescent Substance Abuse, 22*(5), 388-406. doi:10.1080/1067828X.2013.788884

Dozier, M., Peloso, E., Lewis, E., Laurenceau, J.-P., & Levine, S. (2008). Effects of an attachment-based intervention on the cortisol production of infants and toddlers in foster care. *Development and Psychopathology, 20*(3), 845-859. doi:10.1017/S0954579408000400

Enders, C. K. (2010). *Applied Missing Data Analysis*. New York: The Guilford Press.

Fan, A. Y. (2015). The methodology flaws in Hinman's acupuncture clinical trial, Part II: Zelen design and effectiveness dilutions. *Journal of Integrative Medicine, 13*(3), 136-139. doi:10.1016/S2095-4964(15)60172-8

Forgatch, M. S., & DeGarmo, D. S. (1999). Parenting through change: An effective prevention program for single mothers. *Journal of Consulting and Clinical Psychology, 67*, 711-724.

Forgatch, M. S., & Domenech Rodríguez, M. M. (2016). Interrupting Coercion: The Iterative Loops Among Theory, Science, and Practice. In T. J. Dishion & J. J. Snyder (Eds.), *The Oxford Handbook of Coercive Relationship Dynamics* (pp. 194). New York, NY: Oxford University Press.

Forgatch, M. S., & Gewirtz, A. H. (2017). The evolution of the Oregon model of Parent Management Training: An intervention for antisocial behavior in children and adolescents. In J. R. Weisz & A. E. Kazdin (Eds.), *Evidence-based Psychotherapies for Children and Adolescents* (pp. 85-102).

Forgatch, M. S., & Patterson, G. R. (2010). Parent Management Training—Oregon Model: An intervention for antisocial behavior in children and adolescents. In J. R. Weisz & A. E. Kazdin (Eds.), *Evidence-based psychotherapies for children and adolescents* (Vol. 2, pp. 159-178). New York, NY: Guilford Press.

Gresham, F. M., & Elliot, S. (1990). Social skills rating scale. *Circle Pines, MN: American Guidance Service*.

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap) - A meta data-driven methodology and workflow process for providing translated research informatics support. *Journal of Biomedical Informatics, 42*(2), 377-381. doi:10.1016/j.jbi.2008.08.010

Hinman, R. S., McCrory, P., Pirotta, M., Relf, I., Forbes, A., Crossley, K. M., . . . Metcalf, B. R. (2014). Acupuncture for chronic knee pain: A randomized clinical trial. *JAMA, 312*(13), 1313-1322. doi:10.1001/jama.2014.12660

Hodges, K., & Wong, M. M. (1996). Psychometric characteristics of a multidimensional measure to assess impairment: The Child and Adolescent Functional Assessment Scale. *Journal of Child and Family Studies, 5*(4), 445-467. doi:10.1007/BF02233865

Hodges, K., Xue, Y., & Wotring, J. (2004). Use of the CAFAS to evaluate outcome for youths with severe emotional disturbance served by public mental health. *Journal of Child and Family Studies, 13*(3), 325-339. doi:10.1023/B:JCFS.0000022038.62940.a3

Horwitz, S. M., Chamberlain, P., Landsverk, J., & Mullican, C. (2010). Improving the mental health of children in child welfare through the implementation of evidence-based parenting interventions. *Administration and Policy in Mental Health, 37*, 27-39. doi:10.1007/s10488-010-0274-3

Howard, W. J., Rhemtulla, M., & Little, T. D. (2015). Using Principal Components as Auxiliary Variables in Missing Data Estimation. *Multivariate Behavioral Research, 50*(3), 285-299. doi:10.1080/00273171.2014.999267

Hurlburt, M. S., Nguyen, K., Reid, J., Webster-Stratton, C., & Zhang, J. (2013). Efficacy of the Incredible Years group parent program with families in Head Start who self-reported a history of child maltreatment. *Child Abuse & Neglect, 37*(8), 531-543. doi:http://dx.doi.org/10.1016/j.chiabu.2012.10.008

Landsverk, J. A., Burns, B. J., Stambaugh, L. F., & Reutz, J. A. R. (2009). Psychosocial interventions for children and adolescents in foster care: Review of research literature. *Child Welfare, 88*(1), 49-69.

Lang, K. M., & Little, T. D. (2018). Principled missing data treatments. *Prevention Science, 19*(3), 284-294. doi:10.1007/s11121-016-0644-5

Lang, K. M., Little, T. D., & Team, P. D. (2018). PcAux: Automatic extraction of auxiliary variables from large datasets [R package]. In.

Leslie, L. K., Hurlburt, M. S., Landsverk, J., Barth, R. P., & Slymen, D. J. (2004). Outpatient mental health services for children in foster care: A national perspective. *Child Abuse & Neglect, 28*(6), 697-712. doi:http://dx.doi.org/10.1016/j.chiabu.2004.01.004

Leve, L. D., Harold, G. T., Chamberlain, P., Landsverk, J. A., Fisher, P. A., & Vostanis, P. (2012). Practitioner review: Children in foster care - vulnerabilities and evidence-based interventions that promote resilience processes. *Journal of Child Psychology and Psychiatry, 53*(2), 1197-1211. doi:10.1111/j.1469-7610.2012.02594.x

Linares, L. O., Montalto, D., Li, M., & Oza, V. S. (2006). A promising parenting intervention in foster care. *Journal of Consulting and Clinical Psychology, 74*(1), 32-41. doi:http://dx.doi.org/10.1037/0022-006X.74.1.32

Lind, T., Bernard, K., Ross, E., & Dozier, M. (2014). Intervention effects on negative affect of CPS-referred children: Results of a randomized clinical trial. *Child Abuse & Neglect, 38*(9), 1459-1467. doi:https://doi.org/10.1016/j.chiabu.2014.04.004

Little, R. J. A., & Rubin, D. B. (2014). *Statistical Analysis with Missing Data*. Hoboken, NJ: John Wiley & Sons.

Mersky, J. P., Topitzes, J., Grant-Savela, S. D., Brondino, M. J., & McNeil, C. B. (2016). Adapting Parent–Child Interaction Therapy to foster care: Outcomes from a randomized trial. *Research on Social Work Practice, 26*(2), 157-167. doi:10.1177/1049731514543023

Oxford, M. L., Marcenko, M., Fleming, C. B., Lohr, M. J., & Spieker, S. J. (2016). Promoting birth parents' relationships with their toddlers upon reunification: Results from Promoting First Relationships® home visiting program. *Children and Youth Services Review, 61*, 109-116. doi:http://dx.doi.org/10.1016/j.childyouth.2015.12.004

Park, J. M., & Ryan, J. P. (2009). Placement and permanency outcomes for children in out of home care by prior inpatient mental health treatment. *Research on Social Work Practice, 19*(1), 42-51. doi:10.1177/1049731508317276

Patterson, G. R. (1982). *A Social Learning Approach: Coercive Family Process*. Eugene, OR: Castalia Publishing Company.

Patterson, G. R., & Chamberlain, P. (1992). A functional analysis of resistance: A neobehavioral perspective. In H. Arkowitz (Ed.), *Why don't people change? New perspectives on resistance and noncompliance* (pp. 220-242). New York, NY: Guilford Press.

Patterson, G. R., Forgatch, M. S., & DeGarmo, D. S. (2010). Cascading effects following intervention. *Development and Psychopathology, 22*(04), 949-970. doi:10.1017/S0954579410000568

Pecora, P. J., Jensen, P. S., Romanelli, L. H., Jackson, L. J., & Ortiz, A. (2009). Mental Health Services for Children Placed in Foster Care: An Overview of Current Challenges. *Child Welfare, 88*(1), 5-26.

Permanency Innovations Initiative Training and Technical Assistance Project (PII-TTAP), & Permanency Innovations Initiative Evaluation Team (PII-ET). (2013). *The PII Approach: Building Implementation and Evaluation Capacity in Child Welfare*. Washington, DC: U.S. Department of Health and Human Services; Administration for Children and Families, Children's Bureau & Office of Planning, Research and Evaluation.

Price, J. M., Chamberlain, P., Landsverk, J., Reid, J. B., Leve, L. D., & Laurent, H. (2008). Effects of a foster parent training intervention on placement changes of children in foster care. *Child Maltreatment, 13*(1), 64-75.

Reed-Ashcraft, K., Kirk, R. S., & Fraser, M. W. (2001). The reliability and validity of the North Carolina family assessment scale. *Research on Social Work Practice, 11*(4), 503-520.

Romanelli, L. H., Landsverk, J., Levitt, J. M., Leslie, L. K., Hurley, M. M., Bellonci, C., . . . Jensen, P. S. (2009). Best Practices for Mental Health in Child Welfare: Screening, Assessment, and Treatment Guidelines. *Child Welfare, 88*(1), 163-188.

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sones.

Testa, M. F., & Poertner, J. (2010). *Fostering Accountability: Using Evidence to Guide and Improve Child Welfare Policy*. New York: Oxford University Press.

Webster-Stratton, C., & Reid, M. (2010). Adapting The Incredible Years, an evidence-based parenting programme, for families involved in the child welfare system. *Journal of Children's Services, 5*(1), 25-42. doi:doi:10.5042/jcs.2010.0115

Wilson, S. R., Rack, J. J., Shi, X., & Norris, A. M. (2008). Comparing physically abusive, neglectful, and non-maltreating parents during interactions with their children: A meta-

analysis of observational studies. *Child Abuse & Neglect, 32*(9), 897-911. doi:http://dx.doi.org/10.1016/j.chiabu.2008.01.003

Zelen, M. (1990). Randomized consent designs for clinical trials: An update. *Statistics in Medicine, 9*, 645-656. doi:10.1002/sim.4780090611

Figure 1. Participant Flow Chart

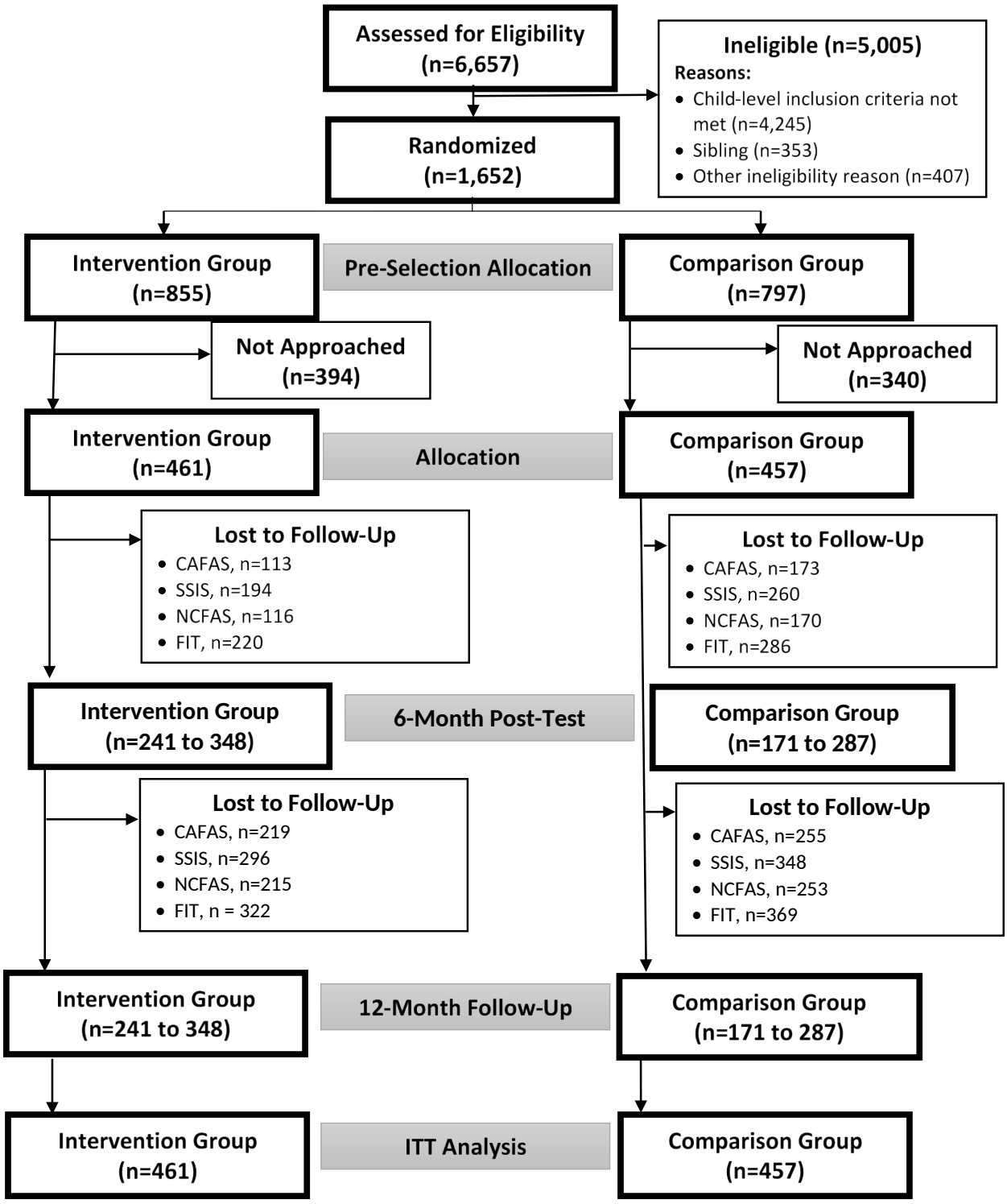

Figure 2

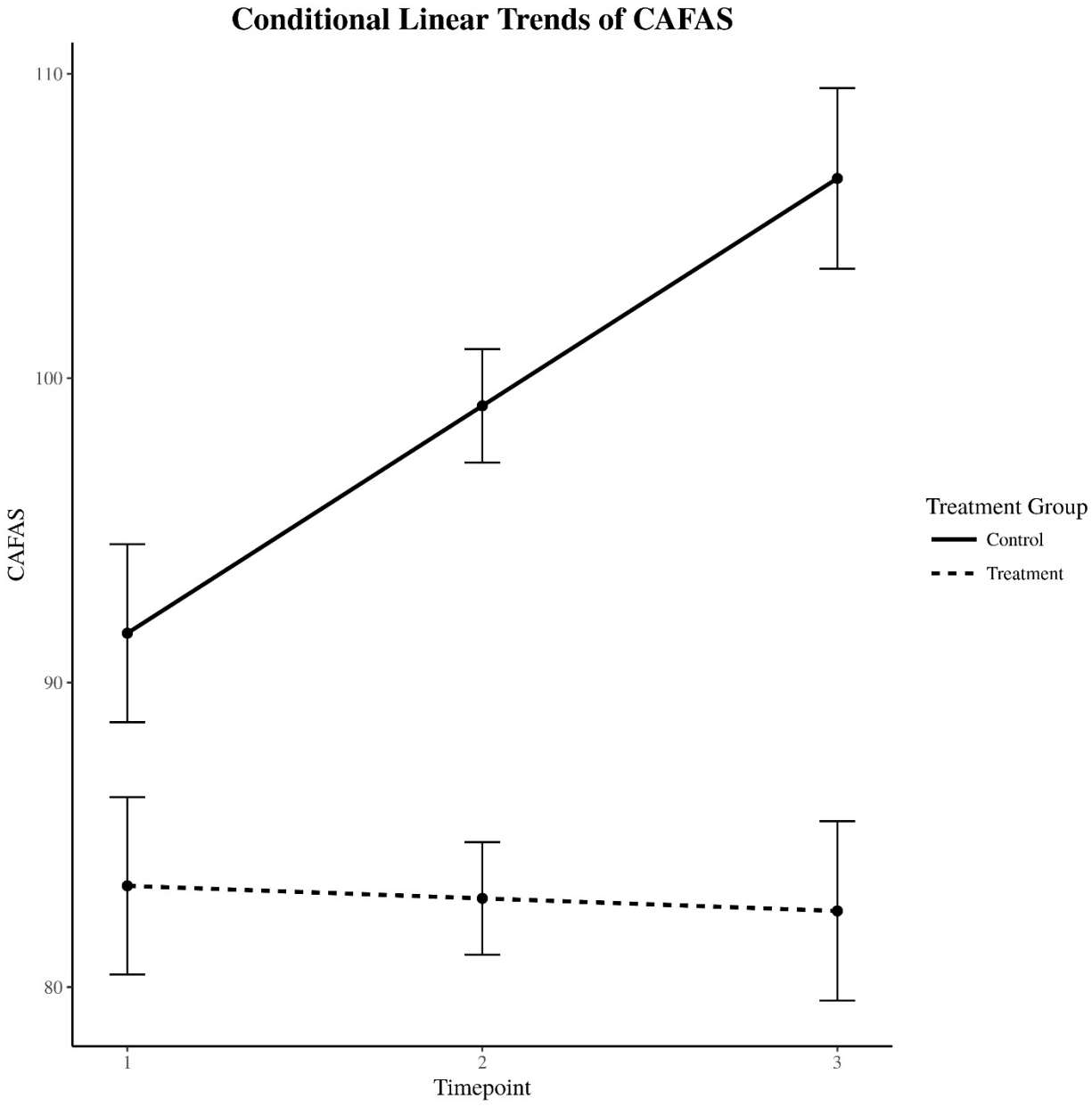
Conditional Linear Trends of CAFAS
110
100
90
80
CAFAS
1
2
3
Timepoint
Treatment Group
Control
Treatment

Figure 3

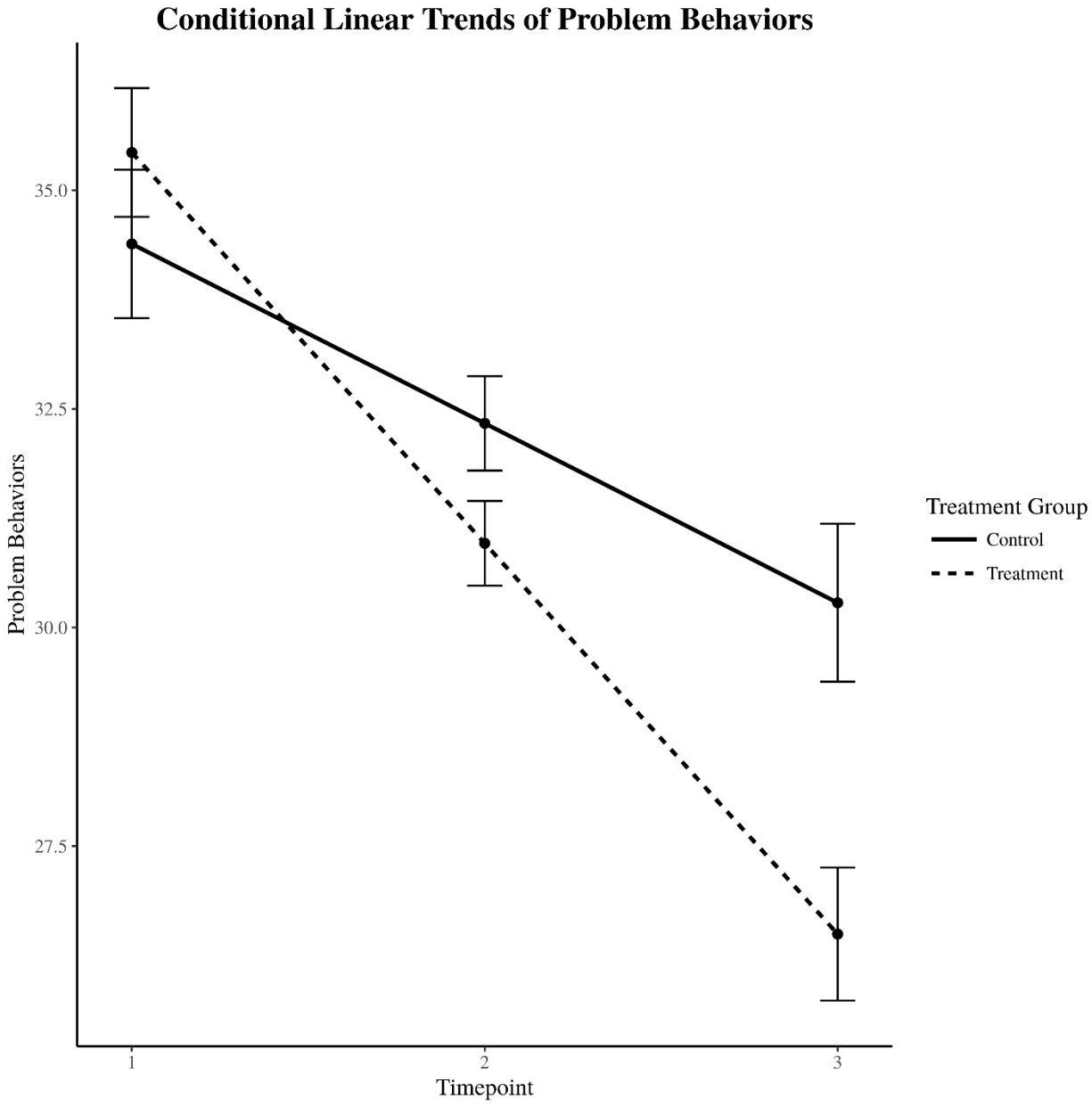
Conditional Linear Trends of Problem Behaviors
Problem Behaviors
35.0
32.5
30.0
27.5
1
2
3
Timepoint
Treatment Group
Control
Treatment

Figure 4

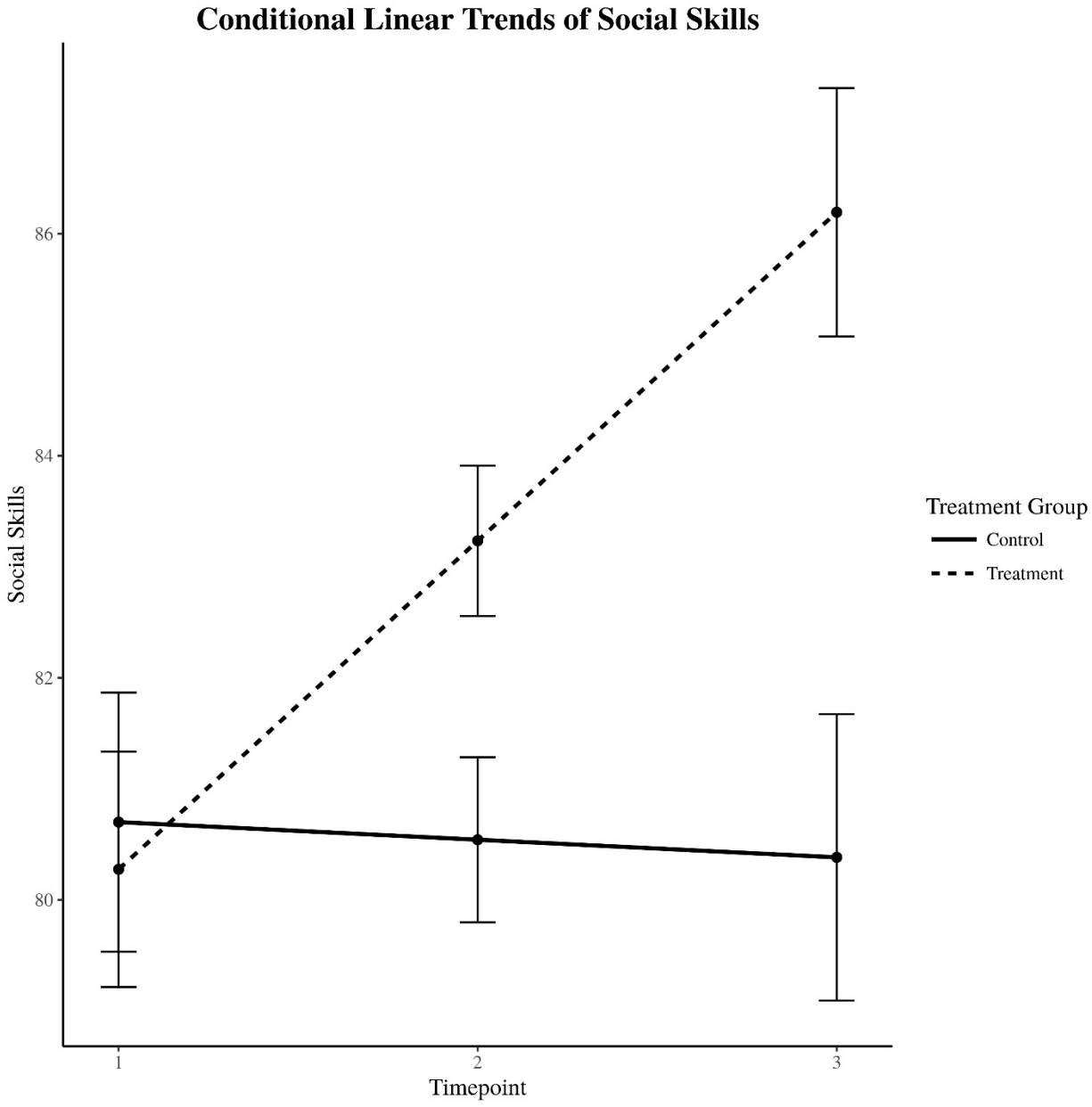
Conditional Linear Trends of Social Skills
Social Skills
86
84
82
80
1
2
3
Timepoint
Treatment Group
Control
Treatment

Figure 5

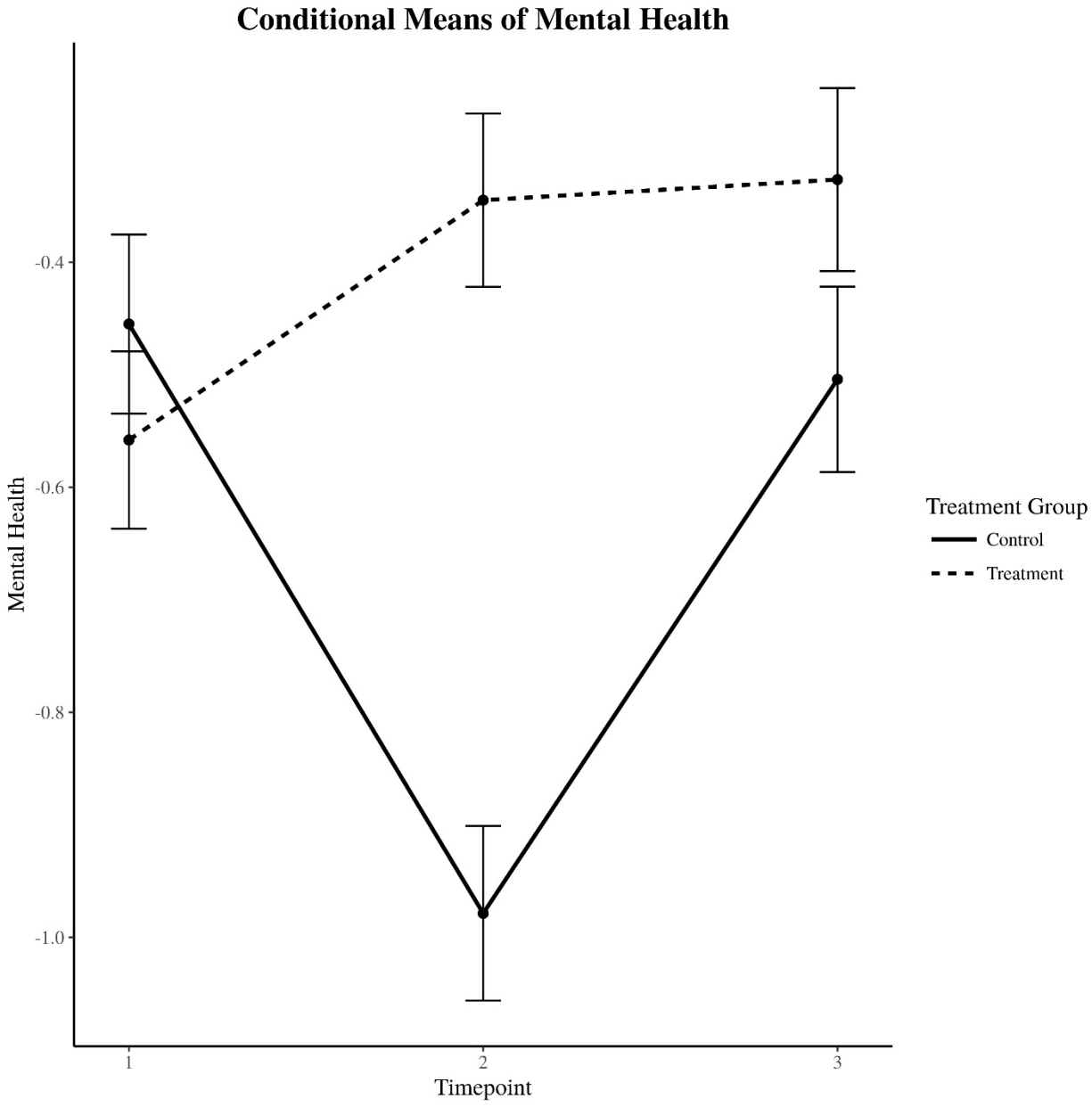
Conditional Means of Mental Health
Mental Health
-0.4
-0.6
-0.8
-1.0
1
2
3
Timepoint
Treatment Group
Control
Treatment

Figure 6

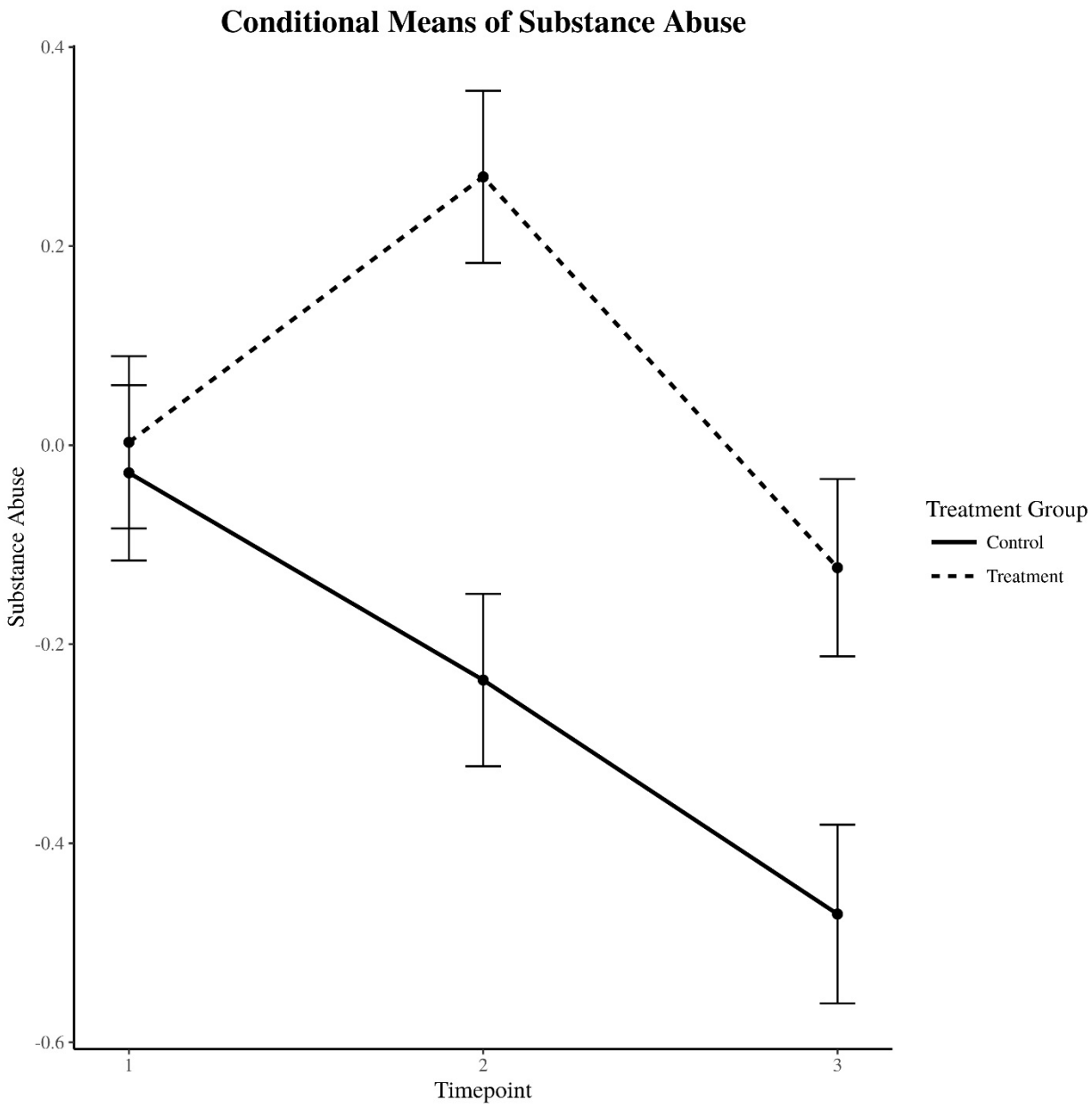
Conditional Means of Substance Abuse
0.4
0.2
0.0
-0.2
-0.4
-0.6
Substance Abuse
1
2
3
Timepoint
Treatment Group
Control
Treatment

Figure 7

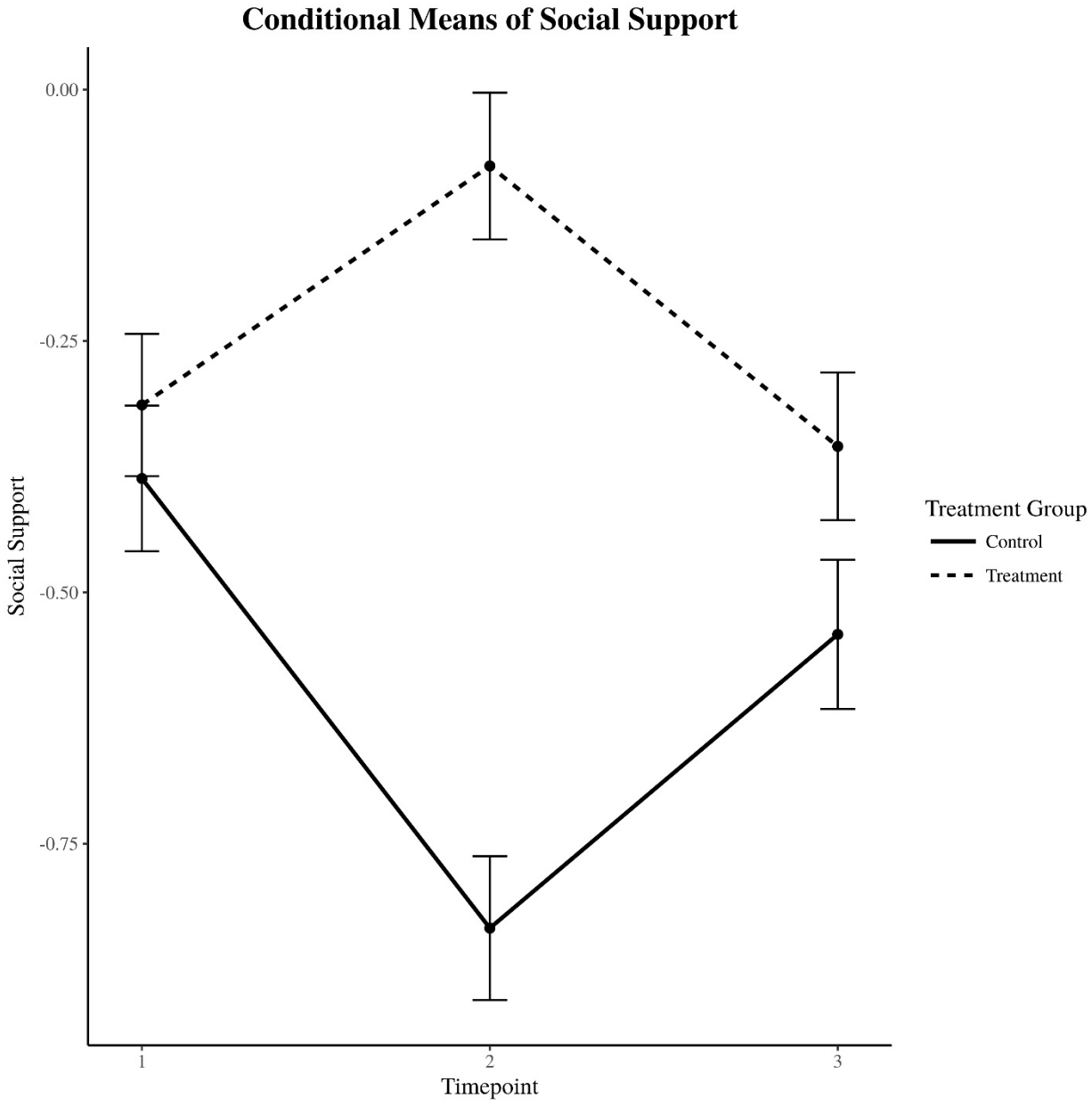
Conditional Means of Social Support
0.00
-0.25
-0.50
-0.75
Social Support
1
2
3
Timepoint
Treatment Group
Control
Treatment

Figure 8

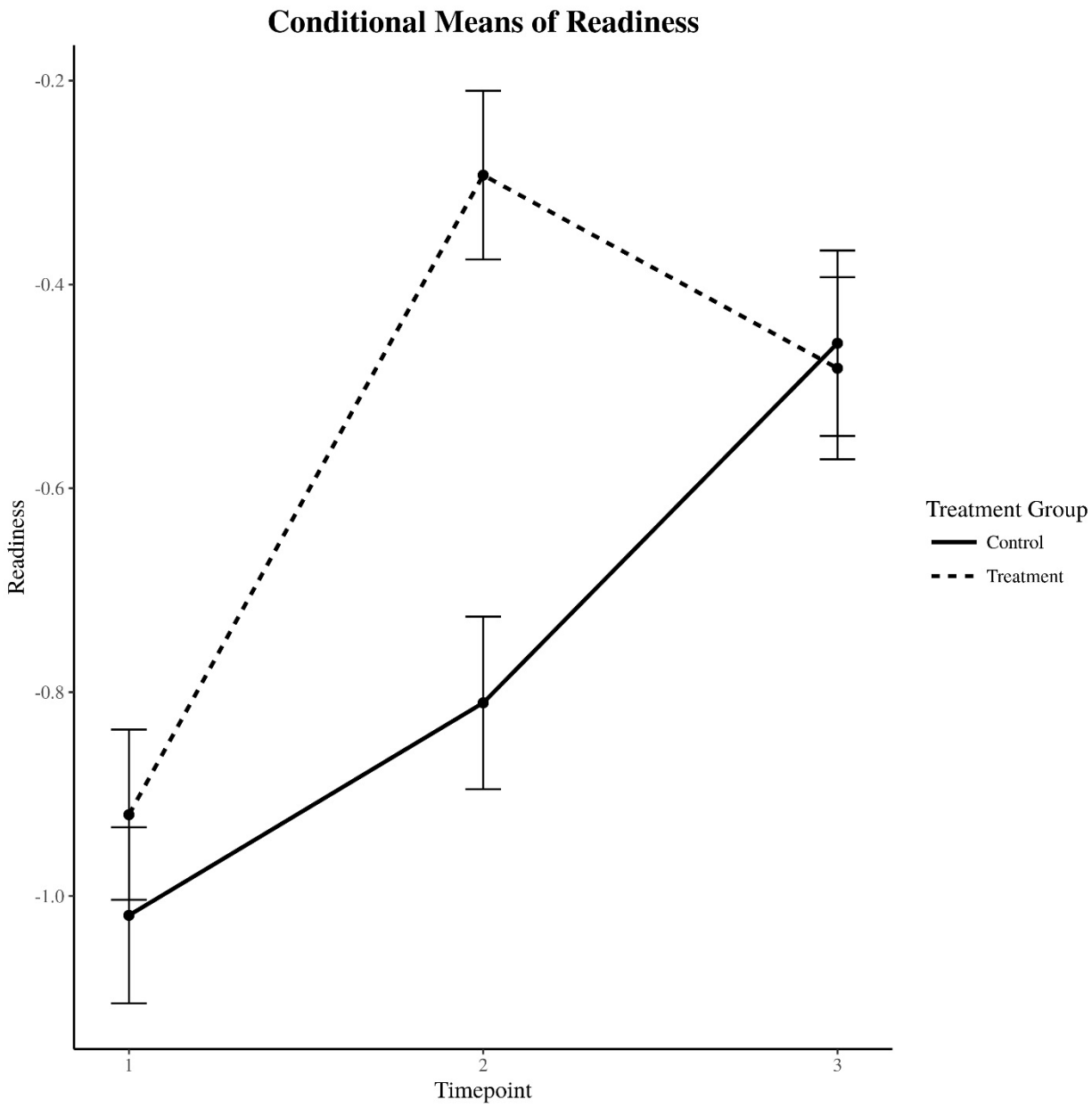
Conditional Means of Readiness
-0.2
-0.4
-0.6
-0.8
-1.0
Readiness
1
2
3
Timepoint
Treatment Group
Control
Treatment

Table 1. Baseline Characteristics of Children and Parents by Study Condition

Characteristic	Total	Intervention Group	Comparison Group	*p*
Child gender is female, *n* (%)	427 (46.5)	204 (44.3)	223 (48.8)	.167
Child race is White, *n* (%)	709 (77.2)	350 (75.9)	359 (78.6)	.341
Child is Latino, *n* (%)	111 (12.3)	55 (11.9)	58 (12.7)	.351
Child age, *M* (*SD*)	11.8 (4.2)	11.6 (4.1)	11.9 (4.3)	.248
Caretaker age at first removal, *M* (*SD*)	38.4 (10.2)	38.2 (10.4)	38.7 (10.1)	.491
Child has diagnosed disability, *n* (%)	493 (53.8)	243 (52.9)	250 (54.7)	.592
Removal reason was physical abuse, *n* (%)	169 (18.4)	87 (18.9)	82 (17.9)	.717
Removal reason was sexual abuse, *n* (%)	57 (6.2)	27 (5.9)	30 (6.6)	.657
Removal reason was neglect, *n* (%)	340 (37.0)	170 (36.9)	170 (37.2)	.919
Removal reason was parent sub abuse, *n* (%)	196 (21.4)	102 (22.1)	94 (20.6)	.565
Parent was single mother, *n* (%)	479 (52.2)	255 (55.3)	224 (49.0)	.056
Parent was single father, *n* (%)	74 (8.1)	38 (8.2)	36 (7.9)	.839
Parents were married/unmarried couple, *n* (%)	365 (39.8)	168 (36.4)	197 (43.1)	.039
Child had prior removals, *n* (%)	197 (21.5)	107 (23.2)	90 (19.7)	.194
Child eligible for IV-E payment, *n* (%)	108 (11.8)	63 (13.7)	45 (9.8)	.073
Time in care at study start, *M* (*SD*)	50.2 (81.0)	54.4 (102)	45.6 (50.8)	.102

Note: Eligible for IV-E payment is a proxy for children's income status. IV-E is the financing mechanism, which provides funds to state agencies and Tribes to assist with the costs of foster care for eligible children. Eligibility is based upon the whether the removal family was eligible for the former Aid to Families with Dependent Children program, effective July 1996.

Table 2. Outcome Means, Standard Deviations, and Cohen's D at Three Time Points

Study Group	Outcome (Measure)	T1		T2		T3		Cohen's D		
		Mean	SD	Mean	SD	Mean	SD	T2 - T1	T3 - T2	T3 - T1
Interv Group	SE Funct (CAFAS)	84.20	41.30	81.40	76.10	83.41	73.56	-0.05	0.03	-0.01
	PB (SSIS)	36.50	15.50	28.80	15.20	27.56	12.82	-0.50	-0.09	-0.63
	SS (SSIS)	79.60	22.00	84.50	22.60	85.54	22.63	0.22	0.05	0.27
	Parenting (FIT)	3.08	0.49	2.89	0.80	3.02	0.91	-0.29	0.15	-0.08
	MH (NCFAS)	-0.56	1.26	-0.34	1.54	-0.33	1.55	0.16	0.01	0.16
	SA (NCFAS)	0.00	1.48	0.27	1.59	-0.12	1.87	0.18	-0.22	-0.07
	Soc Sup (NCFAS)	-0.32	1.48	-0.08	1.74	-0.36	1.50	0.15	-0.17	-0.03
	Ready (NCFAS)	-0.92	1.72	-0.30	1.71	-0.48	1.87	0.36	-0.10	0.24
Control Group	SE Funct (CAFAS)	87.48	40.70	107.40	82.60	102.42	81.44	0.31	-0.06	0.23
	PB (SSIS)	35.15	16.20	30.80	13.90	31.04	13.40	-0.29	0.02	-0.28
	SS (SSIS)	80.62	22.90	80.70	21.60	80.29	22.81	0.00	-0.02	-0.01
	Parenting (FIT)	3.03	0.49	2.90	0.76	2.92	0.90	-0.20	0.02	-0.15
	MH (NCFAS)	-0.45	1.25	-0.98	1.56	-0.5	1.74	-0.37	0.29	-0.03
	SA (NCFAS)	-0.03	1.47	-0.24	1.69	-0.47	1.95	-0.13	-0.13	-0.25
	Soc Sup (NCFAS)	-0.39	1.48	-0.83	1.77	-0.54	1.67	-0.27	0.17	-0.10
	Ready (NCFAS)	-1.03	1.68	-0.81	1.88	-0.46	1.84	0.12	0.19	0.32

Notes: SE Funct = social-emotional functioning; PB = problem behavior; SS = social skills; MH = mental health; SA = substance abuse; Soc Sup = social supports; Ready = readiness for reunification.

Table 3. Models for Children's Outcomes (Tier 1 Results)

Outcome	Unconditional Moderated Model	Unstandardized Estimate	SE	t	p	FMI
Social-emotional functioning (CAFAS)	Intercept	106.56	3.49	30.58	<0.001	0.02
	Time	7.47	2.16	3.47	<0.001	0.02
	Intervention	-24.05	4.70	-5.12	<0.001	0.03
	Intervention*Time	-7.88	2.93	-2.69	0.01	0.02
Problem behaviors (SSIS/PB)	Intercept	30.28	0.88	34.57	<0.001	0.56
	Time	-2.10	0.70	-2.94	0.003	0.51
	Intervention	-3.79	0.98	-3.86	<0.001	0.32
	Intervention*Time	-2.42	0.79	-3.07	0.002	0.26
Social skills (SSIS/SS)	Intercept	80.38	1.29	62.12	<0.001	0.44
	Time	-0.16	0.99	-0.17	0.87	0.42
	Intervention	5.81	1.53	3.79	<0.001	0.13
	Intervention*Time	3.12	1.17	2.67	0.008	0.18

Notes: These results represent Tier 1 analysis in which time was modeled as a linear trend (T1 = -2, T2 = -1, T3 = 0). SE = standard error. FMI = fraction of missing information, which represents the proportion of the total sampling variance that is due to missing data.

Table 4. Models for Caregiver Functioning Outcomes (Tier 2 Results)

Outcome	Unconditional Moderated Model	Unstandardized Estimate	SE	t	p	FMI
Mental Health (NCFAS)	Intercept	-0.46	0.07	-6.15	<.001	0.14
Time 1(Reference)	Time 2	-0.52	0.11	-4.79	<.001	0.09
	Time 3	-0.05	0.11	-0.43	0.67	0.13
	Intervention	-0.10	0.11	-0.99	0.33	0.15
	Intervention*Time 2	0.74	0.15	4.80	<.001	0.10
	Intervention*Time 3	0.28	0.16	1.77	0.08	0.15
Substance Use (NCFAS)	Intercept	-0.03	0.08	-0.34	0.73	0.08
Time 1(Reference)	Time 2	-0.21	0.12	-1.71	0.09	0.06
	Time 3	-0.44	0.13	-3.52	<.001	0.09
	Intervention	0.03	0.12	0.27	0.79	0.08
	Intervention*Time 2	0.47	0.17	2.81	0.005	0.06
	Intervention*Time 3	0.32	0.18	1.80	0.07	0.08
Social Supports (NCFAS)	Intercept	-0.39	0.06	-6.19	<.001	0.13
Time 1(Reference)	Time 2	-0.45	0.10	-4.59	<.001	0.08
	Time 3	-0.16	0.10	-1.50	0.14	0.13
	Intervention	0.07	0.09	0.83	0.41	0.10
	Intervention*Time 2	0.69	0.14	4.99	<.001	0.08
	Intervention*Time 3	0.12	0.14	0.81	0.42	0.11
Readiness (NCFAS)	Intercept	-1.02	0.08	-13.31	<.001	0.18
Time 1(Reference)	Time 2	0.21	0.12	1.79	0.07	0.14
	Time 3	0.56	0.12	4.58	<.001	0.19
	Intervention	0.10	0.11	0.92	0.36	0.16
	Intervention*Time 2	0.42	0.16	2.59	0.01	0.12
	Intervention*Time 3	-0.12	0.17	-0.71	0.48	0.19

Notes: These results represent Tier 2 analysis in which time was dummy coded with T1 as the reference group. Tier 1 results are provided in the supplemental file. SE = standard error. FMI = fraction of missing information, which represents the proportion of the total sampling variance that is due to missing data.

Supplemental File

Table 5. Models for Child Outcomes with Intervention Group Reverse-Coded (Tier 1 Results)

Outcome	Unconditional Moderated Model	Unstandardized Estimate	SE	t	p	FMI
Social-emotional functioning (CAFAS)	Intercept	82.51	3.14	26.26	<0.001	0.02
	Time	-0.41	1.98	-0.21	0.836	0.02
	Intervention	24.05	4.70	5.12	<0.001	0.03
	Intervention*Time	7.88	2.93	2.69	0.01	0.02
Problem behaviors (SSIS/PB)	Intercept	26.49	0.72	36.91	<0.001	0.38
	Time	-4.47	0.56	-7.96	<0.001	0.30
	Intervention	3.79	0.98	3.86	<0.001	0.32
	Intervention*Time	2.42	0.79	3.07	0.002	0.26
Social skills (SSIS/SS)	Intercept	86.19	1.12	76.72	<0.001	0.26
	Time	2.96	0.85	3.50	<0.001	0.24
	Intervention	-5.81	1.53	-3.79	<0.001	0.20
	Intervention*Time	-3.12	1.17	-2.67	0.008	0.18

Notes: Tier 1 analysis modeled time as a linear trend (T1 = -2, T2 = -1, T3 = 0). SE = standard error. FMI = fraction of missing information, which represents the proportion of the total sampling variance that is due to missing data.

Table 6. Models for Parenting Outcome (Tier 1 and Tier 2 Results)

Outcome	Model	Unstandardized Estimate	SE	t	p	FMI
Effective Parenting (FIT)	**Tier 1**					
	Intercept	2.89	0.19	15.16	<0.001	0.96
	Time	-0.06	0.12	-0.47	0.64	0.96
	Intervention	0.07	0.14	0.49	0.63	0.87
	Intervention*Time	0.03	0.09	0.29	0.77	0.85
	Tier 2					
Effective Parenting (FIT)	Intercept	3.03	0.06	50.89	<.001	0.86
Time 1(Reference)	Time 2	-0.13	0.13	-0.99	0.32	0.90
	Time 3	-0.11	0.24	-0.47	0.64	0.96
	Intervention	0.05	0.05	0.91	0.36	0.62
	Intervention*Time 2	-0.06	0.11	-0.53	0.60	0.69
	Intervention*Time 3	0.05	0.18	0.29	0.77	0.85

Notes: Tier 1 analysis modeled time as a linear trend (T1 = -2, T2 = -1, T3 = 0). In Tier 2 analysis, time was dummy coded with T1 as the reference group. SE = standard error. FMI = fraction of missing information, which represents the proportion of the total sampling variance that is due to missing data.

Table 7. Models for Parenting Outcome with Intervention Group Reverse-Coded (Tier 1 and Tier 2 Results)

Outcome	Model	Unstandardized Estimate	SE	t	p	FMI
Effective Parenting (FIT)	**Tier 1**					
	Intercept	2.97	0.12	24.26	<0.001	0.91
	Time	-0.03	0.08	-0.38	0.71	0.90
	Intervention	-0.07	0.14	-0.49	0.63	0.87
	Intervention*Time	-0.03	0.09	-0.29	0.77	0.85
	Tier 2					
Effective Parenting (FIT)	Intercept	3.08	0.05	64.28	<.001	0.77
Time 1(Reference)	Time 2	-0.19	0.12	-1.59	0.11	0.87
	Time 3	-0.06	0.16	-0.38	0.71	0.91
	Intervention	-0.05	0.05	-0.91	0.36	0.62
	Intervention*Time 2	0.06	0.11	0.53	0.59	0.69
	Intervention*Time 3	-0.05	0.18	-0.29	0.77	0.85

Notes: Tier 1 analysis modeled time as a linear trend (T1 = -2, T2 = -1, T3 = 0). In Tier 2 analysis, time was dummy coded with T1 as the reference group. SE = standard error. FMI = fraction of missing information, which represents the proportion of the total sampling variance that is due to missing data.

Table 8. Models for Caregiver Functioning Outcomes (Tier 1 Results)

Outcome	Model	Unstandardized Estimate	SE	t	p	FMI
Mental Health (NCFAS)	Intercept	-0.67	0.08	-8.46	<0.001	0.14
	Time	-0.03	0.06	-0.43	0.67	0.13
	Intervention	0.38	0.11	3.51	<0.001	0.14
	Intervention*Time	0.14	0.08	1.78	0.08	0.15
Substance Use (NCFAS)	Intercept	-0.47	0.09	-5.41	<0.001	0.09
	Time	-0.22	0.06	-3.51	<0.001	0.09
	Intervention	0.45	0.12	3.81	<0.001	0.09
	Intervention*Time	0.16	0.09	1.80	0.07	0.08
Social Supports (NCFAS)	Intercept	-0.67	0.08	-8.95	<0.001	0.10
	Time	-0.08	0.05	-1.50	0.14	0.13
	Intervention	0.40	0.10	3.96	<0.001	0.10
	Intervention*Time	0.06	0.07	0.82	0.41	0.11
Readiness for Reunify (NCFAS)	Intercept	-0.48	0.097	-5.58	<0.001	0.19
	Time	0.28	0.06	4.57	<0.001	0.19
	Intervention	0.14	0.12	1.12	0.26	0.19
	Intervention*Time	-0.06	0.09	-0.70	0.48	0.19

Notes: Tier 1 analysis modeled time as a linear trend (T1 = -2, T2 = -1, T3 = 0). SE = standard error. FMI = fraction of missing information, which represents the proportion of the total sampling variance that is due to missing data.

Table 9. Models for Caregiver Functioning with Intervention Group Reverse-Coded (Tier 2 Results)

Outcome	Unconditional Moderated Model	Unstandardized Estimate	SE	t	p	FMI
Mental Health (NCFAS)	Intercept	-0.56	0.07	-7.58	<.001	0.12
Time 1(Reference)	Time 2	0.21	0.11	2.01	0.04	0.09
	Time 3	0.23	0.11	2.16	0.03	0.14
	Intervention	0.10	0.11	0.99	0.32	0.14
	Intervention*Time 2	-0.74	0.15	-4.80	<.001	0.10
	Intervention*Time 3	-0.28	0.16	-1.78	0.08	0.15
Substance Use (NCFAS)	Intercept	0.003	0.08	0.04	0.97	0.05
Time 1(Reference)	Time 2	0.27	0.12	2.30	0.02	0.05
	Time 3	-0.13	0.12	-1.03	0.30	0.07
	Intervention	-0.03	0.12	-0.26	0.79	0.07
	Intervention*Time 2	-0.48	0.17	-2.82	0.005	0.06
	Intervention*Time 3	-0.32	0.18	-1.81	0.07	0.08
Social Supports (NCFAS)	Intercept	-0.32	0.06	-5.12	<.001	0.08
Time 1(Reference)	Time 2	0.24	0.10	2.47	0.01	0.09
	Time 3	-0.04	0.10	-0.42	0.68	0.10
	Intervention	-0.07	0.09	-0.82	0.41	0.10
	Intervention*Time 2	-0.69	0.14	-4.99	<.001	0.09
	Intervention*Time 3	-0.12	0.14	-0.82	0.41	0.11
Readiness (NCFAS)	Intercept	-0.92	0.07	-12.55	<.001	0.12
Time 1(Reference)	Time 2	0.63	0.11	5.65	<.001	0.08
	Time 3	0.44	0.12	3.61	<.001	0.16
	Intervention	-0.10	0.11	-0.92	0.36	0.16
	Intervention*Time 2	-0.42	0.16	-2.59	0.01	0.12
	Intervention*Time 3	0.12	0.17	0.71	0.48	0.18

Notes: In Tier 2 analysis, time was dummy coded with T1 as the reference group. SE = standard error. FMI = fraction of missing information, which represents the proportion of the total sampling variance that is due to missing data.