American Psychological Association

Journal of Family Psychology

ISSN: 0893-3200

2020, Vol. 34, No. 1, 79–89
http://dx.doi.org/10.1037/fam0000594

Parental Report of Outcomes From a Randomized Trial of In-Home Family Services

Kristin Duppong Hurley and Matthew C. Lambert
University of Nebraska–Lincoln

Irina Patwardhan, Jay L. Ringle, and Ronald W. Thompson
Father Flanagan's Boys Home, Boys Town, Nebraska

Jennifer Farley
University of Nebraska–Lincoln

This study conducted a randomized trial to examine the efficacy of the Boys Town In-Home Family Services (IHFS) program for families of high-risk youth. Participants were recruited from a state helpline for families struggling with poor family functioning and child emotional or behavioral issues. Consent was obtained for 300 of which 152 were randomly assigned to participate in IHFS for 3–4 months and 148 were assigned to the services as usual comparison group. For the families in the treatment group, 18% did not participant in the intervention, and 66% of families received 20 or more service hours. Parent report data were collected at intake, post, as well 6 and 12 months after post data collection. Data were collected on constructs such as caregiver strain, family functioning, parenting, family resources, and parent report of child behavior. Piecewise analyses of the intake to post data indicated significantly greater reductions in caregiver strain for the treatment condition. Given the conservative corrections for the use of multiple tests, no other measures demonstrated significant differences. For the piecewise model of the maintenance phase, there were no significant differences between groups aside from caregiver strain that showed a significant improvement for the comparison condition. Supplementary dose-response analyses indicated that for most families there was an ideal dosage of about 25–75 hr to bring about the largest improvements in caregiver strain, parenting skills, and child behavior.

Keywords: emotional or behavioral needs, in-home services, parent-training, family functioning

Supplemental materials: http://dx.doi.org/10.1037/fam0000594.supp

Raising children can be stressful; especially if the family is experiencing dysfunction and the child has emotional or behavioral challenges. For families experiencing significant distress because of poor family functioning that includes child behavioral or emotional problems, one promising intervention approach is in-home services provided individually to families, typically at the family's home, by a trained provider (e.g., Chaffin, Hecht, Bard, Silovsky, & Beasley, 2012; Fraser, Walton, Lewis, Pecora, & Walton, 1996; Sanders, Kirby, Tellegen, & Day, 2014; Schweitzer, Pecora, Nelson, Walters, & Blythe, 2015). In-home programs

This article was published Online First October 10, 2019.

Kristin Duppong Hurley and Matthew C. Lambert, Department of Special Education and Communication Disorders, University of Nebraska–Lincoln; Irina Patwardhan, Jay L. Ringle, and Ronald W. Thompson, Father Flanagan's Boys Home, Boys Town, Nebraska; Jennifer Farley, Department of Special Education and Communication Disorders, University of Nebraska–Lincoln.

Highlights from the implementation and outcome analyses were shared via presentations to agency staff and the Teaching Family Association annual conference using graphical representations of the outcomes in the fall of 2018.

This research was primarily supported by a research contract from Father Flanagan's Boys Home to the University of Nebraska–Lincoln. This article was supported in part by the U.S. Department of Education through Grant R324B160033 to the University of Nebraska–Lincoln. The opinions expressed are those of the authors and do not necessarily represent the views of Father Flanagan's Boys Home, the Institute of Education Sciences or the U.S. Department of Education. Kristin Duppong Hurley contributed substantially to introduction, discussion, method, interpretation of results, and she was the PI of study. Matthew C. Lambert contributed substantially to results and methods, interpretation of results, and he was the statistician for the study. Irina Patwardhan contributed the methods, tables, and review of manuscript. Jay L. Ringle contributed to the methods, tables and review of manuscript. Ronald W. Thompson contributed to the introduction, interpretation of results, and review of the manuscript. Jennifer Farley contributed to the methods and review of the manuscript. We appreciate the assistance from Lori Synhorst with data analyses and tables, Heidi Menard for assistance with formatting, all of the research staff and students at the Academy for Child and Family Well Being at University of Nebraska–Lincoln for their efforts on this study, the Boys Town In-Home staff including Family Consultants and supervisors, and the wonderful families that participated in the study.

Correspondence concerning this article should be addressed to Kristin Duppong Hurley, Department of Special Education and Communication Disorders, University of Nebraska–Lincoln, Lincoln, NE 68583-0738. E-mail: kristin.hurley@unl.edu

provide intensive family services along with case management and often serve high-risk families in child welfare, juvenile justice, or mental health settings. The common characteristics of in-home programs include families as the unit of focus, the home as the service delivery setting, small caseloads with a team providing 24/7 crises care, individualized services to improve family functioning and parenting, connecting families to formal and informal supports and networks, with services provided every week for a set duration (Schweitzer et al., 2015). In-home programs work to improve outcomes in areas such as caregiver stress, parenting, family functioning, family access to resources, and child behavior (e.g., Chaffin et al., 2012; Lewis, 2005; Sanders et al., 2014) with some programs focused on keeping families intact (e.g., Schweitzer et al., 2015). While outcomes for in-home family interventions when compared with no services are promising (e.g., Chaffin et al., 2012) other studies of in-home services for high-risk families have mixed results (e.g., Schweitzer et al., 2015; Silovsky et al., 2011; Wilson et al., 2012). More high-quality research is needed on in-home programs that serve high-risk families, such as expanding research on widely used programs that lack rigorous efficacy trials (Mason, Fleming, Thompson, Haggerty, & Snyder, 2014). One such intervention that has not been part of a rigorous efficacy trial is the Boys Town In-Home Family Services (IHFS) intervention that is implemented at 11 sites across the United States, serving about 3,500 families in 2017.

IHFS was developed over 30 years ago to serve families with high caregiver strain, poor functioning skills, ineffective parenting strategies, difficulty accessing formal and informal supports, and children with significant emotional and behavioral needs, such as those served by child welfare. The theoretical model of IHFS is an adaptation of the Teaching-Family Model (TFM), which was originally developed as a family style, therapeutic, cognitive–behavioral, residential program for at-risk youth (Phillips, Phillips, Fixsen, & Wolf, 1971). The TFM has promising research evidence according to the National Registry of Evidence-based Programs and Practices and the California Evidence-Based Clearinghouse for Child Welfare and it has been applied successfully to other treatment settings including family foster care, school classroom behavior management, workshop-based parent training, and in-home family intervention. One randomized trial of TFM for in-home services found improvements in child behavior problems, parent–child relationships, and parental provision of physical care and resources (Lewis, 2005).

The IHFS program has a hypothesized theory of change that starts with the provision of individualized, needs-driven services to families with a focus building strong relationships through quick and early solutions. This means that within the first couple of visits to the home the Family Consultant, or in-home service provider, will address a pressing family need to establish strong engagement by reducing parental stress. The Family Consultant works to improve family functioning that should ultimately improve child behavioral and emotional functioning. The primary method for achieving this aim is to coach the family on how to effectively parent their child. Families are also connected to any needed community resources or supports. Several preliminary studies of IHFS have shown positive outcomes for parenting, family functioning, parental stress, and child behavior as well as acceptable model fidelity (Duppong Hurley, Griffith, Casey, Ingram, & Simpson, 2011; Duppong Hurley et al., 2012; Ingram, Cash, Oats, Simpson, & Thompson, 2015; Parra, Ross, Ringle, Samson, & Thompson, 2016). Given the promising evidence, along with demonstrated successful scale-up, the IHFS program is a good candidate for a rigorous efficacy trial.

Our goal was to conduct a randomized study of IHFS with at-risk families following as much as possible the guidelines for transparent randomized studies (e.g., Schulz, Altman, & Moher, 2010). The primary intent-to-treat analyses included the outcomes of caregiver strain, parenting skills, family functioning, family resources, and child emotional and behavioral functioning, similar to prepost studies of IHFS that lacked a comparison group (Duppong Hurley et al., 2012). We hypothesized that participants randomly assigned in the Boys Town IHFS program would demonstrate improved performance over a comparison group at posttest and that gains would be maintained at follow-up. Secondary analyses included moderation analyses of the primary outcomes by frequently used demographic information such as child age and sex, if child is receiving special education services, household income, and single parent household. Many at-risk families also have caregivers with a range of challenging circumstances such as parental depression, substance use, mental health issues, and parenting difficulties (Parra et al., 2016). We created a cumulative family risk variable to examine if outcomes vary by family risk characteristics. The cumulative risk approach addresses the accumulation of the risk factors, rather than the severity of risks or duration of the risk exposure, and is constructed by adding together multiple dichotomous risk affecting family life (Evans, Li, & Whipple, 2013). We assessed the implementation of IHFS in regard to content of core components and dosage. We also conducted dose-response analyses to examine the optimal dose to maximize.

Method

This randomized control trial of the efficacy of the IHFS intervention was conducted from August 2012 until October 2017 with families of children with emotional or behavioral needs. Recruitment was conducted from August 2012 until June 2016. Intervention services were concluded for the final participants by the end of August 2016. Follow-up data collection was concluded by the end of October, 2017. Families and children were randomized to either (a) the IHFS intervention; or (b) a comparison condition of services as usual (SAU) using a random number generator to accomplish a 1:1 allocation ratio with permutated blocks of two participants. Randomization was conducted by the study statistician. Data collection points were conducted at the intake, discharge, 6- and 12-months follow up.

Participants

Families were eligible to participate if they called the family helpline with parenting or child behavior issues, had children ages 5–14, if the caregiver was fluent in English, and the family lived within a local geographic region. Even though families could have called for assistance with multiple children at home, they were asked to identify a target child and provided reports only for this child. It is routine practice for helpline staff to check-back with families several days after the initial family call to ensure that supports are being arranged and that the situation is improving.

During these check-back calls, eligible families were invited to participate in the study. If a family was interested in learning more about the study, they gave permission for their name and phone number to be given to the University research team. Research associates then contacted the family via the phone to provide additional details about the study and obtain informed consent over the phone. Next, families were invited to complete the intake assessment either via paper and pencil or online. After an intake assessment protocol was completed, families were randomly assigned to services as usual or the IHFS intervention. Thus, both conditions were receiving services recommended by the helpline, but those families randomly assigned to the treatment condition also were invited to receive the IHFS services.

A CONSORT diagram of study participation is shown in online supplemental materials Figure 1. Out of 1,262 families that met eligibility requirements, 505 families agreed to talk with researchers about the study, 377 consented to participate, and 300 completed all intake measures. Thus, the final sample was comprised of 300 families. Reasons for nonparticipation included, disappointment with random group assignment, lack of time, and discomfort with home visits. Demographic characteristics of the treatment and SAU groups, which are shown in the online supplemental materials Table 1, were comparable with respect to youth sex ($\chi^2_{(1)} = 0.61$, $p = .437$), youth age ($t_{(298)} = -1.58$, $p = .116$), youth race ($\chi^2_{(1)} = 0.49$, $p = .482$), school identified disability ($\chi^2_{(1)} = 1.76$, $p = .185$), caregiver education ($U = 10{,}236.50$, $z = -1.24$, $p = .215$), and family annual income ($U = 9265.00$, $z = -1.25$, $p = .211$), with no statistically significant differences found between the groups. Preliminary power analyses, based on regression analysis, indicated that a sample size of 300 would allow us to detect a standardized mean difference effect size of 0.27 at posttest when alpha was set at .05 and other predictors in the regression model explained 30% of the variance of the outcome measure.

Procedure

This study was approved by the University Institutional Review Board and all human subject protocols were followed. Consent was collected from a parent or caregiver by University staff. Data were collected by University research staff at baseline (TIME1), and participants were then randomly assigned to either the IHFS intervention or services as usual, followed with posttesting (TIME2) about 3–4 months after the intake assessment. Follow-up data were collected at 6 (Time 3) and 12 months (Time 4) after posttest. Families in the treatment condition were also invited by research staff to participate in implementation and satisfaction measures at 4, 8, and 12 weeks after intake. All data were obtained from parent-report questionnaires that were returned to the University by mail or completed online.

Program dosage. Intervention dosage information (e.g., dates of services, hours of services received) was collected on families assigned to the treatment condition using Boys Town service provision data systems and was then shared with University research staff using secure, shared servers. Variables included Length of Services (in days) as well as Total Service Hours provided during program participation. Intervention service delivery was intended to last about 3–4 months.

IHFS fidelity observation. Treatment families were invited to participate in recordings of their selected sessions with their Family Consultant. Up to three sessions were recorded per family, capturing early, middle, and near the end of services. I-Pad Mini's were used to record each session. While video recordings were made, only the audio content was coded. The recordings were then imported into Nvivo software and coded by research assistants into the core IHFS categories of Engagement/Relationship Building, Family Risk Screen, Social Network Map, Assessment Activities, Parenting Skills, Supports and Resources-Skills, Supports and Resources-Concrete, and Service Planning and Documentation. During the multiyear Study 24 research assistants participated in coding and all reached 97% of reliability during training. Forty-five percent of study recordings were checked for reliability, with 98% indicating acceptable reliability. Any disagreements were discussed and senior research staff was consulted, as needed.

Intervention core components. Boys Town IHFS consists of five program components: family engagement, assessment and service planning, parent and life skill training, assisting with needed resources and supports, and case closure planning (Ingram et al., 2015). There is an emphasis on family engagement in services with shared decision-making and individualized goals. Related to this is the focus of Family Consultants—the trained service providers—to quickly work with families to find a solutions for small practical stressors or problems that the family is currently facing. The intent is to reduce the stress of the family and encourage engagement by showing the ability rapidly improve some aspect of the family's life. Assessment and service planning was conducted with families using a tool called the Strengths and Stressors (Berry, 2009), an adaptation of the North Carolina Family Assessment Scale (Kirk & Reed-Ashcraft, 1998). Assessment and service planning was focused on family goal setting and progress assessment and occurred throughout the intervention. Parent training included specific skill training using an adaptation of the Common Sense Parenting program, which has been rated as Supported by Research by the California Evidence-Based Clearinghouse (Burke, Herron, & Barnes, 2006; Mason et al., 2016). Tools including instruction, modeling, and role play were used to teach skills to address identified child behavior problems and other issues facing the child, family, or both. Throughout the intervention, Family Consultants also helped families build both informal and formal supports essential to maintaining progress after case closure.

All program components are fully manualized, and there is a well-developed system for staff training, supervision, and model fidelity assessment (Ingram et al., 2015). Services are provided by a Family Consultant, who has at least a bachelor's degree in human services, and receives 80 hr of preservice training along with weekly individual supervision and ongoing training and staff development support. Supervision includes observation, coaching, and fidelity checks. Program implementation fidelity and quality are also supported by an agency-wide performance management system that includes electronic dashboards to monitor both model fidelity assessments and outcomes. The agency estimated the hourly cost of IHFS for the study, including supervision and administrative support, at $103.50.

Measures

Caregiver Strain Questionnaire (CGSQ; Brannan, Heflinger, & Bickman, 1997). The CGSQ assess the amount of strain associated with caring for children with behavioral difficulties. It includes seven items that are rated on a 5-point scale ranging from 1 = *not at all a problem* to 5 = *very much a problem*. Higher scores indicate greater caregiver strain. The CGSQ has demonstrated acceptable internal consistency for Objective Strain (four items; $\alpha = .92$), and Subjective Strain (three items; $\alpha = .86$; Brannan et al., 1997). It has also shown to be a valid measure when compared with other measures of parenting strain and distress (Brannan et al., 1997). For this study, total score was used ($\alpha = .88$).

Family Assessment Device (FAD; Epstein, Baldwin, & Bishop, 1983). The FAD assesses family's ability to make decisions regarding their functioning as a family (e.g., "Making decisions is a problem"). It includes 12 items on *general family functioning* that are rated on a 4-point scale ranging from 1 = *strongly disagree* to 4 = *strongly agree* ($\alpha = .91$). Higher scores indicate poorer functioning. Studies support construct and concurrent validity of this measure (Boterhoven de Haan, Hafekost, Lawrence, Sawyer, & Zubrick, 2015; Staccini, Tomba, Grandi, & Keitner, 2015).

Alabama Parenting Questionnaire (APQ; Frick, 1991). The APQ assess parenting practices (i.e., parental involvement) associated with the disruptive behaviors in children, is widely used, and has acceptable reliability ($\alpha = .67$ to $.80$) and validity (Hurley, Huscroft-D'Angelo, Trout, Griffith, & Epstein, 2014; Shelton, Frick, & Wootton, 1996). The APQ includes 42 items that are rated on a 5-point Likert scale ranging from 1 = *never* to 5 = *always*. As no total scores are provided, four subscales were identified for inclusion in the primary analyses: (a) parental involvement with children (10 items, $\alpha = .78$); (b) positive parenting (six items, $\alpha = .84$); (c) poor monitoring, (10 items, $\alpha = .80$); (d) inconsistent discipline (six items, $\alpha = .73$). For the parental involvement and positive parenting subscales, higher scores represent more positive parenting behaviors. For the poor monitoring and inconsistent discipline subscales, higher score represent less positive parenting behaviors.

Parenting Scale (PS; Arnold, O'Leary, Wolff, & Acker, 1993). The PS measures disciplinary practices in parents associated with the externalizing behavior in children and has acceptable reliability and validity (Hurley et al., 2014). Items are phrased as hypothetical situations, and parents are asked to rate how they would react to their child's behavior using a 7-point Likert scale with the endpoints ranging from 1 = *functional* to 7 = *dysfunctional*. The current study used a short 10-item form of the Parenting Scale (Reitman et al., 2001) that includes Overreactivity (five items, $\alpha = .74$) and Laxness (five items, $\alpha = .75$). Research supports the validity of the PS factors via meaningfully strong correlations between the factors and a variety of other measures for both parents (Rhoades & O'Leary, 2007). Higher scores represent less positive parenting behaviors.

Family Resource Scale (FRS; Dunst & Leet, 1987). The FRS assesses specific aspects of perceived family resources and consists of 30 items that are rated on a 5-point Likert scale ranging from 1 = *not at all adequate* to 5 = *almost always adequate* in regard to how well each family need is met. Current analyses supported a four-factor structure of the FRS (Patwardhan, Duppong Hurley, Lambert, & Ringle, 2019): basic needs (10 items, $\alpha = .89$), extra money and time (13 items, $\alpha = .94$), time for family (two items, $\alpha = .90$), and essential care (four items, $\alpha = .64$). Van Horn, Bellis, and Snyder (2001) found this scale to demonstrate good external and convergent validity with other measures of family resources. Higher scores on each of the subscales represent greater needs.

Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997). The SDQ is a 25-item questionnaire designed to assess child behavioral and emotional problems, as well as a Total Difficulties score. A series of studies conducted by Goodman and colleagues indicate strong psychometric properties in both community and clinical samples (Bourdon, Goodman, Rae, Simpson, & Koretz, 2005; Goodman, 2001). For example, parent-reported had acceptable internal reliability ($\alpha = .82$). The SDQ also displays convergent validity in comparison with both clinical judgments (Mathai, Anderson, & Bourne, 2002) and more established questionnaires, such as the Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2001). For this study, the SDQ was completed by parents, with higher scores indicating greater child behavior difficulties. The Total Difficulties subscale was used for analyses ($\alpha = .80$).

Services received. To examine comparability of intervention and SAU group services, families completed a survey at intake, post, 6 and 12 months follow-up that included a series of questions regarding a wide range of services their child "ever" received or received within a specific time-frame (e.g., the past 3 months). The topics of the services received survey were based on the topics covered in the Child and Adolescent Services Assessment (Ascher, Farmer, Burns, & Angold, 1996). A summative variable was created indicating if the child received mental health or social services ever at intake or during a specific time frame on any of 15 service setting items. These 15 items related to emotional or behavioral services included five items related to overnight, out-of-home services (psychiatric hospital, psychiatric ward in a general hospital, detox/drug clinic, residential treatment center, or therapeutic foster home) and 10 items related to outpatient services (day programs, outpatient drug/detox clinic, mental health center, community health center, crisis center, in home counseling/crisis services, independent practice psychologist, school mental health provider, school provided special education services in behavior, or other professional social services).

IHFS Services Received Survey. The IHFS Services Received Survey was developed specifically for the study for parents to identify the degree to which Family Consultants provided services related to the core components of the IHFS intervention, including 12 items covering financial and basic needs, employment, and parenting. The Services Received Survey provides a general description of services provided during the program and was completed at posttest.

Moderating variables. Six variables were used to test whether or not the treatment was differentially effective across groups of youth or caregivers. These variables included: (a) youth age (continuous), (b) youth sex (binary), (c) whether or not the youth had an individualized education program (binary), (d) whether or not the youth lived in a single-parent household (binary), (e) annual income (binary, >$30,000 vs. ≤$30,000), and (f) family risk level (continuous, sum of six risk factors for the

parent). All six potential moderating variables were collected via parent-report at intake into the study. While five of the six potential moderators were basic demographic variables, family risk level was a composite score defined as the sum of six binary risk factors reported by the caregiver (depression, mental illness, substance abuse, criminal behavior, homelessness, or investigated for report of child abuse/neglect).

Statistical Analyses

Effectiveness analyses. Longitudinal data were analyzed using piecewise multilevel modeling (Raudenbush & Bryk, 2012, p. 178–181) implemented in HLM v7 software. Each model was estimated using restricted maximum likelihood so all participants could be included in the analysis even when a participant was missing data on one or more time points. Including all possible participants whether or not full data were available tends to yield less biased (or potentially unbiased) results when data are missing completely at random or missing at random (Allison, 2012).

Piecewise modeling has also been described as "discontinuous growth" modeling (Singer & Willett, 2003) and is particularly useful for evaluating differences between treatment and follow-up phases of an intervention study (Atkins, 2005). The first piecewise segment of time included the baseline and discharge observation points. This segment represented the time when participants were receiving the intervention. The second piecewise segment of time included the follow-up observation points, and represented maintenance of the intervention effects. Initially, unconditional multilevel models (Equation 1) were fit for each outcome to determine if random effects were statistically significant. If Level-2 variances were statistically significant, the random effects were included in the conditional models (Equation 2); conversely, if the variances were nonsignificant, then the random effects were excluded from the conditional models.

$$Y_{ti} = \beta_{00} + \beta_{10} \times \text{TIME1}_{ti} + \beta_{20} \times \text{TIME2}_{ti} + r_{0i} + r_{1i} \times \text{TIME1} + r_{2i} \times \text{TIME2} + e_{ti} \quad (1)$$

$$Y_{ti} = \beta_{00} + \beta_{01} \times \textit{Condition}_i + \beta_{10} \times \text{TIME1}_{ti} + \beta_{11} \times \textit{Condition}_i \times \text{TIME1}_{ti} + \beta_{20} \times \text{TIME2}_{ti} + \beta_{21} \times \textit{Condition}_i \times \text{TIME2}_{ti} + r_{0i} + r_{1i} \times \text{TIME1} + r_{2i} \times \text{TIME2} + e_{ti} \quad (2)$$

In both the unconditional and conditional models the parameters were as follows: β_{00} is the baseline mean for the SAU group, β_{01} is the additive effect on the baseline mean for the treatment group, β_{10} is the mean rate of change for the SAU group during the first piecewise segment of time, β_{11} is the additive effect on the mean rate of change for the treatment group during the first piecewise segment of time, β_{20} is the mean rate of change for the SAU group during the second piecewise segment of time, β_{21} is the additive effect on the mean rate of change for the treatment group during the second piecewise segment of time, r_o is the variance of the baseline score for an individual, r_1 is the variance of the rate of change for the first piecewise segment of time for an individual, r_2 is the variance of the rate of change for the second piecewise segment of time for an individual, and e is the residual for a score a time t for individual *i*.

To test for possible moderated treatment effects, we extended the models described above to include the moderator variable and an interaction term between the treatment indicators and the moderator variable (grand-centered when the moderator was a continuous variable). A statistical significant interaction term indicated that the treatment was differentially effective across levels of the moderator. Because the chance of a making a Type I error was highly inflated (see section below), statistically significant interactions were plotted to further probe the meaningfulness of the moderated effects.

Multiple tests. Twelve tests of main effects of the intervention were analyzed which resulted in Type I error rate inflation. With a total of 12 tests, there was a 46% chance of detecting one statistically significant test ($p < .05$) even if all tests are actually nonsignificant. For test of moderation, six moderators were tested for each of the 12 outcomes resulting in a total of 72 tests. With 72 tests there was a 97.5% chance of detecting at least one statistically significant test even if all tests are actually nonsignificant. For tests of main effects, we recommend interpreting individual tests using a conservative alpha level of .0043 (that sets the family wise Type I error rate at .05; Smolkowski et al., 2017). For moderation analyses, we recommend interpreting individual tests using an alpha level of .001. Tests with probability values less than .05, but greater than the adjusted criterion, might be considered "suggestive" or "promising" (Benjamin et al., 2018).

Effect sizes. Hedges *g* effect size was computed for the first piecewise segment of the multilevel models using the approach suggested by the U.S. Department of Education's National Center for Education Evaluation and Regional Assistance within the Institute of Education Sciences (2014) and Feingold (2009, 2013). This effect size represents the primary impact of the intervention (i.e., standardized differences at posttest) after accounting for individual pretest differences. Separate Hedges *g* effect sizes were computed for the "end of study" differences between participants in the two conditions.

Dose-response analysis. A similar piecewise longitudinal HLM was used to evaluate the associations between dosage and gains made during the intervention for participants in the treatment condition. We hypothesized that the associations would be curvilinear; low and high dosages would be related to weaker effects while moderate dosages would be related to greater effects. Accordingly, we used the total number of service hours and its square to predict the intercept and slopes of the outcome measures over time. The following equation was used to estimate these models:

$$Y_{ti} = \beta_{00} + \beta_{01} \times \textit{HRS}_i + \beta_{02} \times \textit{HRS}_i + \beta_{10} \times \text{TIME1}_{ti} + \beta_{11} \times \textit{HRS}_i \times \text{TIME1}_{ti} + \beta_{12} \times \textit{HRS}_i^2 \times \text{TIME1}_{ti} + \beta_{20} \times \text{TIME2}_{ti} + \beta_{21} \times \textit{HRS}_i \times \text{TIME2}_{ti} + \beta_{22} \times \textit{HRS}_i^2 \times \text{TIME2}_{ti} + r_{0i} + r_{1i} \times \text{TIME1} + r_{2i} \times \text{TIME2} + e_{ti}$$

where β_{11} (linear slope) and β_{12} (quadratic slope) parameters indicate the effect of dose on the slope (i.e., change) from pretest to posttest. However, the statistical significance of the individual parameters was not of primary interest because the goal of these analyses was to describe the practical significance of the associations between dose and gains during the intervention. To this end, we graphed the quadratic functions to visually depict these dose-

response associations, and computed the difference in gains at various points along the quadratic function.

Results

Implementation Data

Dosage. For families randomly assigned to the treatment condition, the length of services ranged from 0 to 213 days, with a mean service length of 104 days (SD = 53.0 days). Total service hours ranged from 0 to 175, with a mean of 35.1 hr (SD = 30.0 hr). Families averaged 1.9 service hours per week, with 18% (n = 27) never engaging in any service hours and 66% (n = 100) of families with more than 20 total service hours.

Services received. Data were also collected on services other than the intervention received by families in the treatment and SAU condition. At intake, 90.8% of families in the treatment condition (n = 138) reported their child received any of 15 emotional or behavioral services in the past compared with 91.9% of families in the SAU condition (n = 136) that represents a statistically nonsignificant difference ($\chi^2_{(1)} = 0.12$, $p = .734$). At posttest, 84.2% of children in the treatment condition received other services during the intervention compared with 80.3% of children in the SAU condition ($\chi^2_{(1)} = 0.14$, $p = .707$). At 6-month follow-up, 80.5% of children in the treatment condition received other services after the intervention compared with 84.4% of children in the control condition ($\chi^2_{(1)} = 0.64$, $p = .425$). Finally, at 12-month follow-up, a statistically significant difference was observed between the families in the two conditions: 69.9% of children in the treatment condition received other services compared with 82.8% of the children in the SAU condition ($\chi^2_{(1)} = 6.05$, $p = .014$).

IHFS fidelity observation. Online supplemental materials Table 2 details the core elements present in 241 recorded meetings. Parenting skills were discussed in nearly all recorded sessions (97.5%) as were engagement and relationship building activities (92.5%). In all, the core components that would be anticipated to be delivered during most sessions were observed with high frequencies. The only item rarely discussed was social network maps (4.2%).

Parent-report of IHFS services received. Parents reported in the IHFS Services Received Survey that Family Consultants were successful in providing the core intervention services to the treatment group, as detailed in online supplemental materials Table 3. Between 70–75% of participants received "quite a bit" or a "great deal" of services related to improving clear and consistent rules, expectations, positive attitudes, and child behavior. Over 50% of participants indicated that services related to budgeting, financial and employment supports, and nutritional services did not apply to them. Only a few families reported not receiving specific services on topics related to budgeting, employment assistance, and financial supports.

Attrition and Baseline Equivalence

Attrition was defined as the 64 participants with baseline data who were missing posttest or follow-up data for at least one observation time point. By the 12 month follow-up observation, we experienced an overall attrition rate of 21.3%; however, the attrition rates were statistically equivalent between the two conditions (24.8 vs. 17.69%) ($\chi^2_{(1)} = 2.3$, $p = .131$). According to the WWC design standards for randomized controlled trials (U.S. Department of Education's National Center for Education Evaluation and Regional Assistance within the Institute of Education Sciences, 2014), the differential attrition rate (7.15 percentage points) observed in this study represents an acceptable level of potential bias when compared with the "optimistic" assumptions underlying the missing data (≤9.9 percentage point difference in rates of attrition; U.S. Department of Education's National Center for Education Evaluation and Regional Assistance within the Institute of Education Sciences, 2014, p. 13). To further evaluate how attrition could have biased results, we tested the sensitivity of baseline scores (across the 12 outcome measures) by using a series of two-way analysis of variances (ANOVAs) where condition, attrition and the interaction between the two factors were predictors of baseline scores. Undesirable sensitivity to attrition would be represented by a statistically significant interaction term indicting that the effect of attrition differs by condition. However, none of the 12 analyses revealed statistically significant interaction effects.

Baseline equivalence for each outcome measure was determined by evaluating the statistical significance and magnitude (as measured by Hedge's g effect size) of the *Condition* effect on the intercept of the multilevel model (β_{01}; difference between control and treatment group means at baseline). Nonsignificant coefficients with effect sizes less than |0.25| suggest equivalence at baseline (U.S. Department of Education's National Center for Education Evaluation and Regional Assistance within the Institute of Education Sciences, 2014). As reported in online supplemental materials Tables 3 and 4, there were no significant Condition effects on the intercept for any of the 12 outcomes, and no effect sizes exceeded |0.25|. The largest effect size difference between conditions at baseline was for the essential care subscale of the FRS (g = 0.200).

Effectiveness

Main effects (TIME1). Descriptive statistics for all four time points for each of the 12 outcomes measures are reported in online supplemental materials Table 4. The main effects of the intervention on the outcome measures were evaluated using piecewise multilevel models (e.g., growth models) where the primary focus for intervention effectiveness was on the statistical significance and the magnitude (e.g., Hedge's g) of the Condition TIME1 interaction (β_{11}). The results for each outcome measure are listed in Tables 1, 2, and 3. Four tests of main effects were statistically significant at the .05 per-test alpha level: (a) Caregiver Strain Questionnaire ($g = -0.402$), (b) Parenting Scale ($g = -0.289$), (c) Family Resource Scale Money & Time subscale (g = 0.200), and (d) Strengths and Difficulties Questionnaire ($g = -0.224$). However, only a single test, CGSQ, was statistically significant at the adjusted alpha criterion of .0043. In the case of the caregiver strain, caregivers in the SAU condition demonstrated a mean change in strain of −0.493 units (β_{10}) while caregivers in the intervention condition demonstrated a mean change of −0.854 units (β_{10} β_{11}). Differences between conditions on "TIME1" slopes resulted in a moderate effect size of $g = -0.402$. In the cases of the other three "suggestive" results, families in the treatment group reported better parenting behaviors, less severe child

Table 1
Results From the Hierarchical Linear Models for Main Effects for Caregiver Strain, Parenting, Family Functioning, and Child Behavior Measures

Effect or statistic			CGSQ	PS	FAD	SDQ
Fixed effects						
Intercept		β_{00}	3.778*** (.075)	3.403*** (.082)	2.188*** (.047)	21.247*** (.561)
Condition		β_{01}	−.089 (.104)	.108 (.115)	−.084 (.066)	−.090 (.784)
TIME1		β_{10}	−.493*** (.088)	−.210* (.083)	−.130** (.044)	−1.535*** (.448)
TIME2		β_{20}	−.175*** (.050)	−.085 (.044)	−.009 (.021)	−.455 (.276)
Condition	TIME1	β_{11}	−.361** (.125)	−.284* (.117)	−.030 (.062)	−1.474* (.632)
Condition	TIME2	β_{21}	.139 (.071)	.018 (.061)	.016 (.030)	−.247 (.392)
Variances						
Intercept		r_0	.345***	.667***	.226***	32.221***
TIME1		r_1	.117	.285**	.066***	—
TIME2		r_2	.048**	.057*	—	1.918***
Residual		e	.470	.301	.100	13.833
p-value						
Condition	TIME1	p	.004	.016	.631	.020
Hedge's *g*						
Posttest		g	−.402	−.289	−.052	−.224
Hedge's *g*						
End of study		g	−.090	−.253	−.004	−.300

Note. CGSQ = Caregiver Strain Questionnaire; PS = Parenting Scale; FAD = Family Assessment Device; SDQ = Strengths and Difficulties Questionnaire.
* $p < .05$. ** $p < .01$. *** $p < .001$.

behavior problems, and greater access to nonessential financial resources.

Maintenance of intervention effects (TIME2). After treatment, the SAU group did not show significant change in the second segment of the piecewise model (β_{20}) nor did the treatment group demonstrate a differential growth rate compared with the SAU group (β_{21}) (Tables 1, 2, and 3) for 11 of the 12 outcomes. Caregiver strain was the exception, which continued to decrease for participants in both the treatment and SAU conditions with caregivers in the SAU condition reporting a slightly greater reduction in strain between posttest and 12-month follow-up observations leading to a substantively smaller "end of study" effect size ($g = -0.090$) than the posttest effect size ($g = -0.402$). Although other outcome measures did not show statistically significant change after the posttest observation, some end of study effect sizes was larger in magnitude than the effect sizes at posttest. For

Table 2
Results From the Hierarchical Linear Models for Main Effects for Additional Parenting Measures

Effect or statistic			APQ_PI	APQ_PP	APQ_PMS	APQ_ID
Fixed effects						
Intercept		β_{00}	36.666*** (.510)	24.438*** (.311)	16.906*** (.556)	15.719*** (.347)
Condition		β_{01}	.362 (.713)	−.073 (.435)	.588 (.773)	−.031 (.486)
TIME1		β_{10}	.937* (.399)	.248 (.286)	.013 (.477)	−1.457*** (.298)
TIME2		β_{20}	.036 (.245)	−.007 (.150)	.043 (.274)	.164 (.164)
Condition	TIME1	β_{11}	.040 (.657)	−.230 (.403)	−.803 (.669)	−.280 (.419)
Condition	TIME2	β_{21}	−.137 (.348)	.129 (.212)	.018 (.387)	−.388 (.230)
Variances						
Intercept		r_0	25.868***	10.29***	29.692***	11.465***
TIME1		r_1	6.168**	3.402***	2.857*	—
TIME2		r_2	1.212*	.600**	2.088***	—
Residual		e	9.936	3.677	10.484	5.760
p-value						
Condition	TIME1	p	.952	.569	.231	.504
Hedge's *g*						
Posttest		g	.007	−.062	−.117	−.065
Hedge's *g*						
End of study		g	−.039	.008	−.123	−.259

Note. APQ_PI = Alabama Parenting Questionnaire: Parent Involvement; APQ_PP = Alabama Parenting Questionnaire: Positive Parenting; APQ_PMS = Alabama Parenting Questionnaire: Poor Monitoring Supervision; APQ_ID = Alabama Parenting Questionnaire: Inconsistent Discipline.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 3
Results From the Hierarchical Linear Models for Main Effects for Family Resource Measures

Effect or statistic			FRS_BN	FRS_MON	FRS_FAM	FRS_CAR
Fixed effects						
Intercept		β_{00}	4.436*** (.053)	2.927*** (.080)	3.847*** (.081)	3.752*** (.111)
Condition		β_{01}	.018 (.075)	.150 (.113)	.158 (.113)	.189 (.155)
TIME1		β_{10}	−.013 (.050)	.072 (.068)	.119 (.089)	.118 (.118)
TIME2		β_{20}	.005 (.028)	.030 (.041)	−.018 (.049)	−.026 (.066)
Condition	TIME1	β_{11}	.125 (.071)	.202* (.097)	−.088 (.125)	.073 (.169)
Condition	TIME2	β_{21}	−.030 (.039)	−.050 (.058)	.082 (.069)	−.064 (.098)
Variances						
Intercept		r_0	.246***	.719***	.422***	.505***
TIME1		r_1	—	.159***	—	—
TIME2		r_2	—	.077***	—	—
Residual		e	.163	.224	.534	.369
p-value						
Condition	TIME1	p	.077	.038	.484	.664
Hedge's *g*						
Posttest		g	.199	.200	−.090	.080
Hedge's *g*						
End of study		g	.098	.105	.078	−.057

Note. FRS_BN = Family Resource Scale: Basic Needs; FRS_MON = Family Resource Scale: Extra Money; FRS_FAM = Family Resource Scale: Family Needs; FRS_CAR = Family Resource Scale: Care Needs.
* $p < .05$. ** $p < .01$. *** $p < .001$.

example, the effect size for the distal outcome of child behavior was larger at the end of the study ($g = -0.300$) than the effect size at posttest. Likewise, the end of study effect size for the Inconsistent Discipline subscale score ($g = -0.259$) was larger than the effect size at posttest.

Moderated intervention effects. Because the primary focus of the main effect tests was on the Condition TIME1 interaction (β_{11}), we tested the moderation of that parameter by six different moderators (described above). Six interactions were statistically significant at the .05 per-test alpha level: (a) APQ Positive Parenting subscale IEP*, (b) APQ Positive Parenting subscale Income, (c) Parenting Scale Income, (d) CQSG IEP, (e) SDQ Risk, and (f) APQ Poor Monitoring Risk*. Two interactions were significant at the adjusted criterion level (denoted by *). Those interactions were further probed using graphical approaches. Neither interaction appeared meaningful in terms of differential growth between baseline and posttest observations (β_{11}) for the two outcomes (APQ Positive Parenting, APQ Poor Monitoring).

Dose-Response Analyses

For the dose-response analyses, the four outcomes that demonstrated statistically significant or suggestive effects were probed further to evaluate the associations between dose and the gains made during the intervention for individuals assigned to the treatment condition. For these analyses, one participant with 175 hr of services was omitted to reduce bias introduced by this outlier value. The numerical results of these analyses are reported in Table 5 in the online supplemental materials. Model-estimated functions describing the associations between dosage and gains during the intervention were plotted in Figure 1.

For the Family Resource Scale (Panel D), there was a weak negative linear association between dosage and gains during the intervention. However, for the Caregiver Strain Questionnaire (Panel A), Parenting Scale (Panel B), and SDQ (Panel C) outcome measures, a clear positive quadratic association (i.e., a U-shaped function) was estimated, where low dosage (≲20 hr) and high dosage (≳20 hr) were related to weaker gains during the intervention, and moderate dosages (25–75 hr) were related to greater gains during the intervention. At different points along the quadratic function, these differences in gains were noteworthy.

The model for caregiver strain predicted that a participant who received 50 hr of services would have a decrease in strain that was 32% greater than a participant who received 10 hr of services and 53% greater than a participant who received 100 hr of services, after adjusting for baseline differences in strain. Based on the statistical model, 51 hr of services was optimal for maximizing gains for caregiver strain. For child behavior severity, the model predicted that the child of a participant who received 50 hr of services would have a decrease in behavior that was 35% greater than the child of a participant who received 10 hr of services and 21% greater than the child of a participant who received 100 hr of services, after adjusting for baseline differences in child behavior severity. Based on the statistical model, 59 hr of services was optimal for maximizing gains for child behavior severity. For parenting practices, the model predicted that the a participant who received 50 hr of services would have a decrease in poor parenting practices that was 146% greater than a participant who received 10 hr of services and 7% greater than a participant who received 100 hr of services, after adjusting for baseline differences in parenting practices. Based on the statistical model, 71 hr of services was optimal for maximizing gains for parenting practices.

Discussion

The focus of this study was to use an intent-to-treat randomized control study to examine the posttest and maintenance effects of the Boys Town IHFS program over services as usual for families of youth with emotional and behavioral needs. Recorded observations of sessions with the families found strong support that core

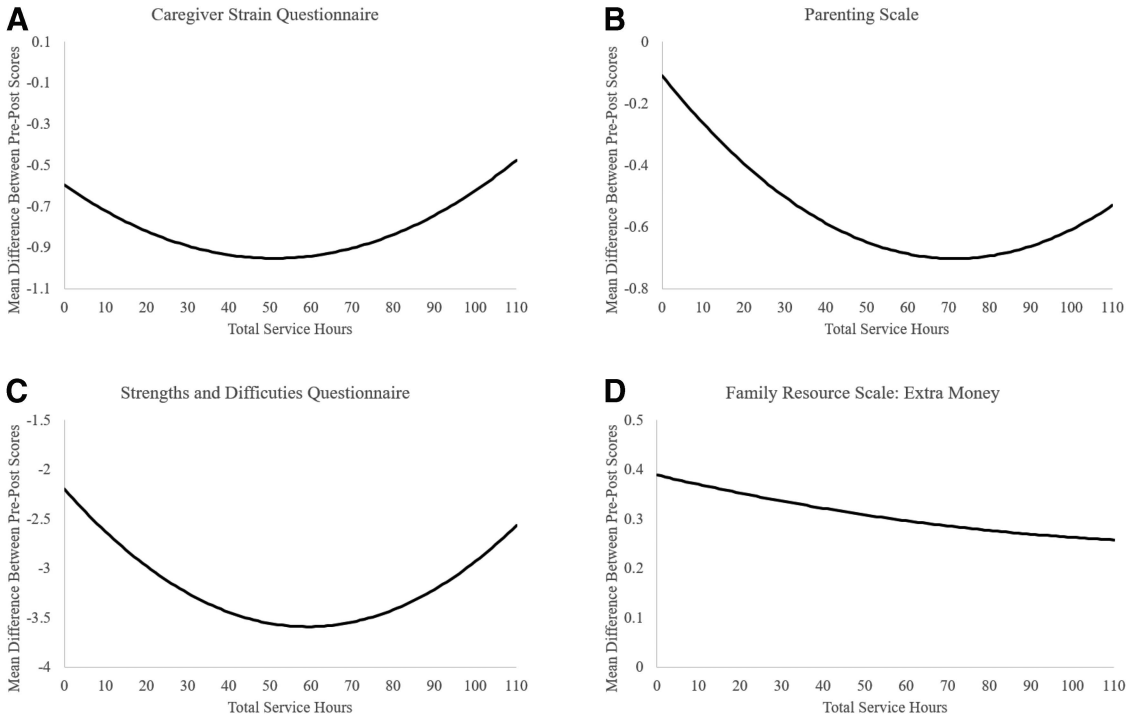

Figure 1. Graphs for dose-response analyses.

components of the IHFS model were being implemented, with parenting skills and engagement/relationship-building activities being observed during nearly every coded session. A few activities had lower reported frequencies, but those activities, such as assessments, were not intended to occur on every visit. The use of social network maps was quite low; which indicates that this aspect of the intervention was infrequent in implementation during routine sessions. However, this may be in part because of Family Consultants typically developing the social network maps very early in services during sessions that were not frequently recorded. Similar to the recorded observations of in-home sessions, parent reports of services received indicated high-rates of endorsement by families for activities such as improving child behavior, promoting positive attitudes, discussing clear and consistent family rules, and family expectations. The activities that were reported less frequently (e.g., services related to budgeting, employment, and nutritional needs) were likely relevant to a smaller percentage of families individualized service plans, which would mean they would be infrequently addressed, if at all, for many families.

One of the challenges with parent-focused programs is engaging families to enroll in and fully participate in the intervention (e.g., Chacko et al., 2016; Ingoldsby, 2010). With regard to dosage, the mean length of services was within the model recommended 3–4 month range with an expected average provision of almost 2 hr of services per week. A sizable portion of families (18%) did not engage in the IHFS program, but overall participation rates were comparable to other parent training interventions (Chacko et al., 2016).

Experimental studies of high-risk families typically compare treatment to services as usual or an alternative approach. For this study services as usual included a considerable amount of support. Specifically, looking at families that reported receiving mental health or social services for their child the rates showed no significant differences between conditions at intake (about 90% of children), post and 6 month follow-up (about 82% of children). The comparison group reported significantly more services used between the 6 and 12 month follow-up (83% SAU vs. 70% treatment). A substantial percentage of children across both groups were receiving additional supports related to their emotional or behavioral health needs, as found in other studies of high-risk families (e.g., Chaffin et al., 2012) that could help explain positive gains on outcomes for participants in the control condition.

The results indicated strong support for the effect of IHFS on caregiver strain at posttest, with an effect size of .402. The Parenting Scale, Family Resources Extra Money and Time subscale, and SDQ each indicated "suggestive" effects, significant at the .05 α but not at the adjusted alpha level. The growth rates for both the treatment and SAU groups were similar in the maintenance phase, aside from caregiver strain, where participants in the SAU group had slightly higher rates of strain reduction compared with the treatment group. The outcome of child behavior demonstrated an even larger effect size at follow-up than at post, as measured by the SDQ. The effect size for this study for parent-report of child behavioral outcomes is similar to the reported adjusted meta-analyses results for similar interventions (Sanders et al., 2014). In all, the findings that there was a significant reduction in caregiver strain at post, with promising improvements in parenting, family resources related to extra time and money, and improved child behavior at posttest and follow-up is encouraging of the potential of the IHFS program. This is especially promising given the considerable amount of reported emotional and behavioral services that were received by families in both study conditions, and the considerable percentage of families in the treatment condition that never engaged in IHFS.

It is intriguing that for measures of caregiver strain, parenting, and child emotional or behavioral problems a quadratic association was found with neither very low or very high doses of services being associated with the strongest gains in outcomes. Thus, for most families there was an ideal dosage of about 25–75 hr to bring about the largest improvements. This makes some intuitive sense, as families that did not engage in services are unlikely to make as significant of improvements. Also, with families with extensive needs and considerable service hours may have had more issues than could be addressed during a time-limited intervention. Alternatively, the family resources measure had a linear relationship where additional dosage was weakly associated with worse outcomes. This finding is puzzling, but may be related to the items in this construct concerning parents finding more time for themselves that could be related to initial efforts to reduce caregiver stress resulting in more personal time.

Limitations

In routine practice, IHFS is most often provided to families referred by child welfare agencies. Therefore, it would be ideal if an effectiveness study of IHFS could be conducted with a child welfare population. However, such studies are ethically challenging to undertake, given the difficulty for families to be able to freely choose to both participate and withdraw from research when required to participate in services by family social service agencies. While not a child welfare population, participants were having enough issues to call a helpline for support. Moreover, about one-third of the youth experienced an out-of-home stay related to their emotional or behavioral needs, indicating high-risk families. Another limitation was that only parental self-report data were collected; thus, the data collected were not "blind" as the parents knew the condition they received. The study also did not collect cost-effectiveness information. It would be helpful if future research could include expensive observational data of family functioning and child behavior that were beyond the financial and logistical scope of this study as well as detailed information on costs of services for treatment and SAU conditions.

Implications

This study demonstrated that IHFS significantly reduced caregiver strain at post assessment for caregivers of youth aged 5–16 that called a statewide helpline for support. It also has suggestive promise for improving parenting skills, family resources for extra money and time, as well as child behavioral functioning at home. However, the results were mixed with no differences found for several intermediate outcomes. Given the costs, in time and money, of in-home services it is highly recommended that additional research be initiated to explore the effectiveness of the approach especially in regard to better understanding the optimal dosage to achieve the largest sustained outcomes. With so many families in need of support to improve parenting skills, family functioning and child emotional and behavioral needs, innovative in-home services research continues to be needed to understand how to best meet the individualized needs of high-risk families.

References

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms and profiles*. Burlington: Research Center for Children, Youth, & Families, University of Vermont.

Allison, P. D. (2012). *Handling missing data by maximum likelihood*. Cary, NC: SAS Institute.

Arnold, D. S., O'Leary, S. G., Wolff, L. S., & Acker, M. M. (1993). The Parenting Scale: A measure of dysfunctional parenting in discipline situations. *Psychological Assessment, 5,* 137–144. http://dx.doi.org/10.1037/1040-3590.5.2.137

Ascher, B. H., Farmer, E. M. Z., Burns, B. J., & Angold, A. (1996). The Child and Adolescent Services Assessment (CASA): Description and psychometrics. *Journal of Emotional and Behavioral Disorders, 4,* 12–20. http://dx.doi.org/10.1177/106342669600400102

Atkins, D. C. (2005). Using multilevel models to analyze couple and family treatment data: Basic and advanced issues. *Journal of Family Psychology, 19,* 98–110. http://dx.doi.org/10.1037/0893-3200.19.1.98

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., . . . Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour, 2,* 6–10. http://dx.doi.org/10.1038/s41562-017-0189-z

Berry, M. (2009). *Strengths and stressors, v. 2.0*. Lawrence, KS: Author.

Boterhoven de Haan, K. L., Hafekost, J., Lawrence, D., Sawyer, M. G., & Zubrick, S. R. (2015). Reliability and validity of a short version of the general functioning subscale of the McMaster family assessment device. *Family Process, 54,* 116–123. http://dx.doi.org/10.1111/famp.12113

Bourdon, K. H., Goodman, R., Rae, D. S., Simpson, G., & Koretz, D. S. (2005). The Strengths and Difficulties Questionnaire: U.S. normative data and psychometric properties. *Child & Adolescent Psychiatry, 44,* 557–564. http://dx.doi.org/10.1097/01.chi.0000159157.57075.c8

Brannan, A. M., Heflinger, C. A., & Bickman, L. (1997). The Caregiver Strain Questionnaire: Measuring the impact on the family of living with a child with serious emotional disturbance. *Journal of Emotional and Behavioral Disorders, 5,* 212–222. http://dx.doi.org/10.1177/106342669700500404

Burke, R., Herron, R., & Barnes, B. A. (2006). *Common sense parenting: Using your head as well as your heart to raise school-aged children*. Boys Town, NE: Boys Town Press.

Chacko, A., Jensen, S. A., Lowry, L. S., Cornwell, M., Chimklis, A., Chan, E., . . . Pulgarin, B. (2016). Engagement in behavioral parent training: Review of the literature and implications for practice. *Clinical Child and Family Psychology Review, 19,* 204–215. http://dx.doi.org/10.1007/s10567-016-0205-2

Chaffin, M., Hecht, D., Bard, D., Silovsky, J. F., & Beasley, W. H. (2012). A statewide trial of the SafeCare home-based services model with parents in Child Protective Services. *Pediatrics, 129,* 509–515. http://dx.doi.org/10.1542/peds.2011-1840

Dunst, C. J., & Leet, H. E. (1987). Measuring the adequacy of resources in households with young children. *Child: Care, Health and Development, 13,* 111–125. http://dx.doi.org/10.1111/j.1365-2214.1987.tb00528.x

Duppong Hurley, K., Griffith, A. K., Casey, K. J., Ingram, S., & Simpson, A. (2011). Behavioral and emotional outcomes of an in-home parent training intervention for young children. *Journal of At-Risk Issues, 16,* 1–7.

Duppong Hurley, K., Griffith, A., Ingram, S., Bolivar, C., Mason, W. A., & Trout, A. (2012). An approach to examining the proximal and intermediate outcomes of an intensive family preservation program. *Journal of Child and Family Studies, 21,* 1003–1017. http://dx.doi.org/10.1007/s10826-011-9563-z

Epstein, N. B., Baldwin, L. M., & Bishop, D. S. (1983). The McMaster family assessment device. *Journal of Marital and Family Therapy, 9,* 171–180. http://dx.doi.org/10.1111/j.1752-0606.1983.tb01497.x

Evans, G. W., Li, D., & Whipple, S. S. (2013). Cumulative risk and child development. *Psychological Bulletin, 139,* 1342–1396. http://dx.doi.org/10.1037/a0031808

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

Feingold, A. (2013). A regression framework for effect size assessments in longitudinal modeling of group differences. *Review of General Psychology, 17,* 111–121. http://dx.doi.org/10.1037/a0030048

Fraser, M. W., Walton, E., Lewis, R. E., Pecora, P. J., & Walton, W. K. (1996). An experiment in family reunification: Correlates of outcomes at one-year follow-up. *Children and Youth Services Review, 18,* 335–361. http://dx.doi.org/10.1016/0190-7409(96)00009-6

Frick, P. J. (1991). *The Alabama Parenting Questionnaire*. Unpublished instrument, Developmental Psychopathology Laboratory, University of New Orleans, New Orleans, LA.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38,* 581–586. http://dx.doi.org/10.1111/j.1469-7610.1997.tb01545.x

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

Hurley, K. D., Huscroft-D'Angelo, J., Trout, A., Griffith, A., & Epstein, M. (2014). Assessing parenting skills and attitudes: A review of the psychometrics of parenting Measures. *Journal of Child and Family Studies, 23,* 812–823. http://dx.doi.org/10.1007/s10826-013-9733-2

Ingoldsby, E. M. (2010). Review of interventions to improve family engagement and retention in parent and child mental health programs. *Journal of Child and Family Studies, 19,* 629–645. http://dx.doi.org/10.1007/s10826-009-9350-2

Ingram, S. D., Cash, S. J., Oats, R. G., Simpson, A., & Thompson, R. W. (2015). Development of an evidence-informed in-home family services model for families and children at risk of abuse and neglect. *Child & Family Social Work, 20,* 139–148. http://dx.doi.org/10.1111/cfs.12061

Kirk, R. S., & Reed-Ashcraft, K. (1998). *User's guide for the North Carolina Family Assessment Scale, Version 2.0*. Jordan Institute for Families, University of North Carolina School of Social Work, Chapel Hill, NC.

Lewis, R. E. (2005). The effectiveness of Families First services: An experimental study. *Children and Youth Services Review, 27,* 499–509. http://dx.doi.org/10.1016/j.childyouth.2004.10.009

Mason, W. A., Fleming, C. B., Thompson, R. W., Haggerty, K. P., & Snyder, J. J. (2014). A framework for testing and promoting expanded dissemination of promising preventive interventions that are being implemented in community settings. *Prevention Science, 15,* 674–683. http://dx.doi.org/10.1007/s11121-013-0409-3

Mason, W. A., January, S.-A. A., Fleming, C. B., Thompson, R. W., Parra, G. R., Haggerty, K. P., & Snyder, J. J. (2016). Parent training to reduce problem behaviors over the transition to high school: Tests of indirect effects through improved emotion regulation skills. *Children and Youth Services Review, 61,* 176–183. http://dx.doi.org/10.1016/j.childyouth.2015.12.022

Mathai, J., Anderson, J., & Bourne, A. (2002). The Strengths and Difficulties Questionnaire (SDQ) as a screening measure prior to admission to a Child and Adolescent Mental Health Service (CAMHS). *Australian e-Journal for the Advancement of Mental Health, 1,* 235–246. http://dx.doi.org/10.5172/jamh.1.3.235

Parra, G. R., Ross, J. R., Ringle, J. L., Samson, N. R., & Thompson, R. W. (2016). Evaluation of Boys Town In-Home Family Services with families referred by child welfare. *Journal of Evidence-Informed Social Work, 13,* 401–411. http://dx.doi.org/10.1080/23761407.2015.1086715

Patwardhan, I., Duppong Hurley, K., Lambert, M., & Ringle, J. A. (2019). The family resource scale: Psychometric properties and validation in a sample of high risk families. *Journal of Psychoeducational Assessment, 37,* 372–381. http://dx.doi.org/10.1177/0734282918769486

Phillips, E. L., Phillips, E. A., Fixsen, D. L., & Wolf, M. M. (1971). Achievement place: Modification of the behaviors of pre-delinquent boys within a token economy. *Journal of Applied Behavior Analysis, 4,* 45–59. http://dx.doi.org/10.1901/jaba.1971.4-45

Raudenbush, S. W., & Bryk, A. S. (2012). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Reitman, D., Currier, R. O., Hupp, S. D., Rhode, P. C., Murphy, M. A., & O'Callaghan, P. M. (2001). Psychometric characteristics of the Parenting Scale in a head start population. *Journal of Clinical Child Psychology, 30,* 514–524. http://dx.doi.org/10.1207/S15374424JCCP3004_08

Rhoades, K. A., & O'Leary, S. G. (2007). Factor structure and validity of the parenting scale. *Journal of Clinical Child and Adolescent Psychology, 36,* 137–146. http://dx.doi.org/10.1080/15374410701274157

Sanders, M. R., Kirby, J. N., Tellegen, C. L., & Day, J. J. (2014). The Triple P-Positive Parenting Program: A systematic review and meta-analysis of a multi-level system of parenting support. *Clinical Psychology Review, 34,* 337–357. http://dx.doi.org/10.1016/j.cpr.2014.04.003

Schulz, K. F., Altman, D. G., & Moher, D., & the CONSORT Group. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *British Medical Journal, 340,* c332. http://dx.doi.org/10.1136/bmj.c332

Schweitzer, D. D., Pecora, P. J., Nelson, K., Walters, B., & Blythe, B. J. (2015). Building the evidence base for intensive family preservation services. *Journal of Public Child Welfare, 9,* 423–443. http://dx.doi.org/10.1080/15548732.2015.1090363

Shelton, K. K., Frick, P. J., & Wootton, J. (1996). Assessment of parenting practices in families of elementary school-age children. *Journal of Clinical Child Psychology, 25,* 317–329. http://dx.doi.org/10.1207/s15374424jccp2503_8

Silovsky, J. F., Bard, D., Chaffin, M., Hecht, D., Burris, L., Owora, A., . . . Lutzker, J. (2011). Prevention of child maltreatment in high-risk rural families: A randomized clinical trial with child welfare outcomes. *Children and Youth Services Review, 33,* 1435–1444. http://dx.doi.org/10.1016/j.childyouth.2011.04.023

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195152968.001.0001

Smolkowski, K., Seeley, J. R., Gau, J. M., Dishion, T. J., Stormshak, E. A., Moore, K. J., . . . Garbacz, S. A. (2017). Effectiveness evaluation of the Positive Family Support intervention: A three-tiered public health delivery model for middle schools. *Journal of School Psychology, 62,* 103–125. http://dx.doi.org/10.1016/j.jsp.2017.03.004

Staccini, L., Tomba, E., Grandi, S., & Keitner, G. I. (2015). The evaluation of family functioning by the family assessment device: A systematic review of studies in adult clinical populations. *Family Process, 54,* 94–115. http://dx.doi.org/10.1111/famp.12098

U.S. Department of Education's National Center for Education Evaluation and Regional Assistance within the Institute of Education Sciences. (2014). *What Works Clearinghouse Procedures and Standards Handbook Version 3.0*. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v3_0_standards_handbook.pdf

Van Horn, M. L., Bellis, J. M., & Snyder, S. W. (2001). Family Resource Scale-Revised: Psychometrics and validation of a measure of family resources in a sample of low-income families. *Journal of Psychoeducational Assessment, 19,* 54–68. http://dx.doi.org/10.1177/073428290101900104

Wilson, P., Rush, R., Hussey, S., Puckering, C., Sim, F., Allely, C. S., . . . Gillberg, C. (2012). How evidence-based is an 'evidence-based parenting program'? A PRISMA systematic review and meta-analysis of Triple P. *BMC Medicine, 10,* 130. http://dx.doi.org/10.1186/1741-7015-10-130

Received July 5, 2018
Revision received July 2, 2019
Accepted August 20, 2019 ■