TRUE Dads:

Including co-parents in a fatherhood intervention

Final Impact Evaluation Report for It's My Community Initiative, Inc.

Oklahoma City, OK

November 19, 2020

Prepared by

Philip A. Cowan, Carolyn Pape Cowan, and Peter Gillette,

University of California, Berkeley

According to the Paperwork Reduction Act of 1995, an agency may not conduct or sponsor, and a person is not required to respond to, a collection of information unless it displays a valid OMB control number. The valid OMB control number for this collection is 0970-0356; this number is valid through 6/30/2021. Public reporting burden for this collection of information is estimated to average 30 hours, including the time for reviewing instructions, gathering and maintaining the data needed, reviewing the collection of information, and revising it. This collection of information is voluntary for individuals, but the information is required from Grantees.

Recommended Citation:

Cowan, P. A., Cowan, C. P., & Gillette, P. (2020). TRUE Dads: Including co-parents in a fatherhood intervention. Berkeley, CA: University of California.

Acknowledgements:

We would like to thank the Administration for Child and Family's Office of Planning, Research, and Evaluation (OPRE) for their support of this project, as well as Mathematica for their consultation and guidance. Many thanks to the staff at It's My Community Initiative (IMCI) for their devoted work in managing and executing the TRUE Dads Intervention, and to the evaluators from the University of Denver for their ongoing collaboration.

Disclosure:

No conflict of interest for any of the authors

This publication was prepared under Grant Number 90FK008302 from the Office of Family Assistance (OFA) within the Administration for Children and Families (ACF), U.S. Department of Health & Human Services (HHS). The views expressed in this report are those of the authors and do not necessarily represent the views or policies of HHS, ACF, or OFA.

Structured Abstract: TRUE Dads: Including co-parents in a fatherhood intervention

Objective. The TRUE Dads fatherhood program focuses on establishing and strengthening three of men's key roles in the family: 1) as fathers, building and maintaining positive engagement with their children; 2) as partners, having and keeping a positive relationship with their co-parenting partners (wife, intimate partner, or other co-parent); and 3) as providers, through fostering employment and economic self-sufficiency. Evaluators used a randomized controlled trial (RCT) design to determine the impacts of the TRUE Dads program on key outcomes in these domains for the intervention group as compared to those in a business-as-usual control group.

Study design. Low-income dads raising a young child (birth to 12 years) and their co-parenting partners were recruited for this study. A total of 1,042 co-parenting teams were randomly assigned to the TRUE Dads group intervention (60%) or a no-treatment control group (40%). Surveys to measure impacts were obtained before random assignment and again one year following enrollment. The intervention included 6 core workshops (18 hours) followed by optional participation in an additional 6 workshops (18 hours) on parenting or couple relationships. All fathers in the intervention condition were offered a chance to participate in WorkForward, a service of Work Ready Oklahoma.

Results. At least one partner completed the one-year follow-up in 360 (85.5%) co-parenting teams in the control group and 531 (85.5%) teams in the TRUE Dads intervention group. One year after entering the study, TRUE Dads program participants showed fewer symptoms of depression, anxiety, and anger, less destructive couple communication and domestic violence, and increased economic self-sufficiency. Supplementary analyses in the TRUE Dads group suggested that reductions in parents' negative symptoms and destructive couple communication led to reductions in negative parenting qualities and children's behavior problems, and to increases in fathers' hours of employment, income, and positive attitudes about employment.

Conclusion. TRUE Dads, an innovative program that works with both parents to strengthen fathers' positive roles in the family, had positive impacts on fathers' personal distress, their involvement in their family relationships, and their economic self-sufficiency. Supplementary analyses suggested that the intervention's impact on personal distress and co-parenting relationship quality might also be associated with reductions in both parents' harsh parenting and the frequency of behavior problems in their children. The TRUE Dads intervention findings demonstrate clearly that the inclusion of a co-parent in a fatherhood intervention and a focus on improving the quality of their relationship are important keys to enhancing responsible fatherhood and children's development. From the perspective of family policy, the results suggest that integration rather than separation of healthy marriage and responsible fatherhood programs could provide synergistic power to programs designed to strengthen low-income parents and their children.

Contents

Tables

Figures

Impact Evaluation of True Dads It's My Community Initiative, Oklahoma City, OK

I. Introduction

A. Introduction and study overview

Decades of research demonstrates that the quality of the relationship a parent shares with a child is fundamental to the well-being of both parent and child (Cabrera, Volling, & Barr, 2018). While most of this research focuses on mothers, fathers make different but equally important contributions to their children's development, (Cowan, Cowan, Cohen, Pruett, & Pruett, 2008). It is also the case that even the most motivated fathers face barriers to engaging with their children. In Oklahoma, many low-income men work in jobs that pay at or near the minimum wage of $7.25 an hour (Bureau of Labor Statistics, 2018). Working long and unpredictable hours on weekdays and weekends, the conventional characteristics of contemporary low-wage work, takes low-income fathers out of the home, leaving them little time for their families and children (Gennetian & Cabrera, 2018). Without additional support and assistance, it can be difficult for fathers to find a pathway to economic self-sufficiency.

Despite the fact that most biological fathers are in a romantic relationship with the mother when their child is born, many fade from their children's lives in subsequent years (Carlson, McLanahan, & Brooks-Gunn, 2009). The importance of this fact lies in the finding that the relationship fathers have with their co-parents is a primary factor in the quality of their relationship with their children. In 2020, in Oklahoma City, 38% of children were living in single-parent families, and fathers were most often the absent parent (Foundation, 2020). When parenting pairs live together as couples, the quality of their relationship is associated with two protective factors that the intervention to be described in this report addresses directly: fathers' ability to co-parent cooperatively and their positive engagement with their child (Achatz & MacAllum, 1994).

Poverty and psychological distress are intimately connected. Economic insecurity and family distress are triggers of parents' depression and anxiety. Oklahoma has one of the highest rates of major depression in adults in the nation (News on 6, 2020), with many needs going unaddressed (Oklahoma, 2020), and this fact has consequences for children because parents' personal distress creates a barrier to positive involvement with their children, to well-regulated relationships between the co-parents, and to active participation in the labor force.

From the brief presentation of these research findings, we can see that low-income fathers in Oklahoma City are at greater risk for father absence, anxiety, depression, unemployment, and distress in their relationships with the co-parents with whom they share responsibility for raising their children. There is an urgent need to provide services to enhance men's positive engagement in family life and improve relationships with both their co-parenting partners and their children, but at the time this study began no such services were available in Oklahoma City or the surrounding counties.

Family strengthening interventions have typically been addressed in separate silos offering couple relationship or father engagement services (Cowan & Cowan, 2018). Existing couple relationship interventions rarely focus systematically on the co-parenting or father-child relationships, and, until very recently, rarely provide employment-related resources. There is extensive research showing that these roles and domains are interconnected. For example, high unresolved conflict between parents is associated with harsh or permissive parenting styles and aggression and/or depression in their children (Cummings & Davies, 2010; Harold, Acquah, Sellers, & Chowdry, 2016). Parental depression has strong links with couple relationship quality (Whisman & Schonbrun, 2010), parenting quality, and children's well-being. Intervention studies indicate that the domains may be causally linked so that change in one family domain produces change in others. For example, a recent RCT of Supporting Father Involvement -- a father engagement intervention using a couples group approach -- showed that over 18 months the intervention significantly reduced conflict between the parents, which, in turn, led to less harsh or permissive parenting styles in fathers' and mothers' parenting strategies, which was associated with fewer acting out/aggressive or depressed/withdrawn behaviors in their children (Cowan, Cowan, Pruett, & Pruett, 2019).

Fatherhood interventions rarely include the father's co-parenting partner. Most are offered to groups of men with male group leaders, yet studies show that the single best predictor of a father's involvement with his child is the quality of his relationship with the child's mother (Carlson, Pilkauskas, McLanahan, & Brooks-Gunn, 2011). Recently, an international review of hundreds of couple, fatherhood, and family-based intervention studies concluded that the few fatherhood interventions that offered group services for both partners together led to positive increases in the quantity and quality of fathers' involvement with their children *and* in the quality of the co-parenting relationship (Panter-Brick et al., 2015). This correlational finding was supported by the results of four rigorous studies including two RCTs, with an intervention approach similar to the one that is the subject of the present report; each intervention offered 16 weekly group meetings to low-income fathers and their co-parents, with a curriculum that covered multiple domains of family functioning. All four studies demonstrated positive effects of the Supporting Father Involvement intervention on individual distress, couple relationship quality, parenting quality, and children's behavior problems (see Cowan & Cowan, 2018 for a summary).

This report describes the implementation and impact of the TRUE Dads program in Oklahoma City. Designed to address the gaps in much of the previous literature on fatherhood interventions, it provides a 16-week curriculum in a group intervention format that integrates a focus on

fathering, parent-child relationships, co-parenting, three-generational relationships, and employment services, with the overarching goal of strengthening fathers' key roles and positive involvement in the family. Co-parenting teams were eligible for the program if their youngest child was between birth and 12 years of age.

The remainder of this introductory section describes the research questions that motivated our research. The Section II describes the program as it was intended to function, including the intervention curriculum, the control condition, and the implementation of the study. Section III describes the study design and data collection procedures for the implementation and impact analyses. Section IV describes the analysis methods and outcome measures. Section V describes the study findings and the estimation approach. Finally, Section VI provides a summary and discussion of the study findings.

B. Primary research questions

Six primary research questions focus on potential impacts of the TRUE Dads program:

1. What is the impact of TRUE Dads relative to a control group on parents' level of depression one year after they began the intervention?
2. What is the impact of TRUE Dads relative to a control group on couple conflict one year after they began the intervention?
3. What is the impact of TRUE Dads relative to a control group on domestic violence one year after they began the intervention?
4. What is the impact of TRUE Dads relative to a control group on father's and co-parent's harsh parenting one year after they began the intervention?
5. What is the impact of TRUE Dads relative to a control group on child behavior problems one year after parents began the intervention?
6. What is the impact of TRUE Dads relative to a control group on fathers' level of employment one year after they began the intervention?

This study and its outcome measures are listed on the registry of clinicaltrials.gov (NCT03069898).

C. Additional research questions

This evaluation provides us with an opportunity to learn more about the mechanisms by which TRUE Dads produces positive impacts on the family. We used structural equation modeling to explore the hypothesis that when participants in TRUE Dads reduce their symptoms of depression and their relationship as co-parents improves (e.g., less conflict, more collaboration), (a) fathers and co-parents will interact with their children more effectively in ways that reduce children's behavior problems, and (b) fathers will be more motivated and organized to obtain employment, retain their jobs, and earn higher wages.

II. Intervention and counterfactual conditions

This section provides a brief description of the curriculum components of the TRUE Dads intervention, the staffing of the TRUE Dads program, and the flow of participants through the program. Table II.1 and the text below describe each intervention component and the business-as-usual counterfactual condition.

A. Description of program as intended

1. Intended components

TRUE Dads services are offered at the It's My Community Initiative (IMCI) offices in downtown Oklahoma City. Program core services are primarily delivered through the TRUE Dads On My Shoulders (TD-OMS) curriculum, which is a series of 6 workshops (18 hours total) adapted from the existing evidence-based intervention On My Shoulders. A father and co-parent—referred to here as a co-parenting team—attend the program together to promote a shared learning experience to help them align expectations, learn skills for more effective communication, and create a stronger commitment to more collaborative shared parenting. The number of co-parenting teams assigned to each workshop cohort varied, with the program's priority placed on starting co-parents in groups as soon as possible after intake.

In addition to the TD-OMS core curriculum (18 hours), TRUE Dads participants could take one of two optional intensive tracks (6 sessions, 18 hours total per track), one focused on couple relationships, the other on parenting. Both members of the co-parenting team must have participated in the TD-OMS core workshops in order for the team to enroll in the relationship intensive. A father could choose to attend the parenting intensive alone, regardless of whether or not his co-parent attended TD-OMS. A co-parent, however, could not attend a track separately from the participating father. In most cases, the co-parenting partners jointly decided which, if any, intensive track to take.

In addition, all *fathers* assigned to the intervention were offered a chance to participate in a two-week (up to 60 hours) intensive WorkForward program to bolster their economic self-sufficiency.

Each co-parenting team was assigned to a coach who met with them for an initial case management meeting and then on an as-needed basis. The coaches also served as a liaison between the co-parents and the TRUE Dads group educators. During the first meeting, the program coaches helped co-parenting teams complete an initial assessment to identify their strengths and needs. The coaches subsequently engaged the co-parenting teams to facilitate workshop attendance, reinforce skills learned in workshops, and make referrals to community resources as needed.

2, Intended curriculum content

The curriculum content of TRUE Dads is guided by our family systems theory of change (Cowan & Cowan, 2019) in which 6 domains of risk and protective factors interact to affect individuals and relationships in the nuclear family (see Appendix A): (1) The history of experiences in the father's and co-parent's family of origin; (2) Co-parents' personal distress (e.g., anxiety and depression); (3) The quality of the relationship between the co-parents; (4) Parent-child relationship quality; (5) Children's behavior problems; and (6) Attaining economic self-sufficiency and facing other external stressors. Our evidence-based theory assumes that interventions to strengthen father involvement and to improve the quality of co-parenting relationships must address issues in all 6 domains.

TRUE Dads On My Shoulders (TD-OMS). TRUE Dads core curriculum is a unique interactive co-parenting adaptation of the individually-focused On My Shoulders curriculum (PREP for Individuals, I; 2011). It was delivered in 6 3-hour group sessions designed to equip men with the skills needed for healthy relationships, especially between the co-parenting partners and with their children. PREP staff re-envisioned the OMS curriculum for TRUE Dads to include material from other evidence-based interventions, thus creating a unique approach to a fatherhood intervention that engages father and co-parent. The adapted curriculum is branded TRUE Dads On My Shoulders (TD-OMS). This adaptation requires that a co-parent be present for the group workshops and adds employment/economic stability and parenting topics to fully address three key areas (healthy relationships, parenting, and economic stability) plus other risk/protective domains (personal distress, life stress, social support, intergenerational family patterns) associated with a fathers' ability to succeed in his major family roles. Themes include couple communication, stress, co-parenting, effective discipline, commitment, and gratitude.

Relationship intensive. The optional couple relationship intensive utilizes the evidence-based PREP 8.0 curriculum (Stanley et al., 2017), delivered in 6 3-hour group sessions, and designed for fathers and their parenting partners to improve essential relationship skills—such as communicating effectively; working as a team; solving challenging problems and managing conflict effectively without resorting to violent strategies; and preserving and enhancing love, commitment, and friendship.

Parenting intensive. The optional parenting intensive was designed for TRUE Dads, with the goal of providing additional support for the co-parenting relationship and each parent's relationship with his or her young children. The parenting intensive, delivered in 6 3-hour group sessions, is an adaptation of the Supporting Father Involvement curriculum (Cowan, Cowan, Pruett, Pruett, & Wong, 2009), with additional material from other evidence-based programs. Themes include the importance of fathers to children's development; stages of child development; parenting styles, with a focus on discipline appropriate to the child's stage; effective communication with children; collaboration between co-parents; reflecting on three-generational relationship patterns; and how to cope with individual, relationship, and family stress.

Employment track. Fathers in co-parenting teams were offered an opportunity to participate in the WorkForward two-week long employment track. Participants in WorkForward received an assessment of their education and job skills, engaged in job-readiness training, and were offered job placement assistance with a robust employer network of in-demand, higher wage occupations. An Employment Development Specialist assisted participants in improving work readiness skills, such as developing an effective resume and preparing for job interviews. Participants also had access to an on-site Resource Center with computers, office supplies, and other materials to aid in enhancing job readiness and attainment. WorkForward was offered for up to 60 hours, although not all fathers completed all 60 hours. For example, fathers may have left the program early to pursue an employment opportunity.

3. Planned dosage and implementation schedule

TD-OMS, parenting, and relationship intensives were offered to program participants in 3-hour group sessions, once per week for six weeks. The dosage was 18 hours for TD-OMS and 18 hours for each intensive. WorkForward was offered to fathers in six-hour sessions each weekday for two consecutive weeks, up to a total of 60 hours.

4. Target population

TRUE Dads is intended for low-income fathers, particularly men of color, who: (a) have a child between birth and age 12, with special emphasis on ages 0-5; (b) are able to involve a co-parent in the study; (c) are 18 years or older; and (d) reside in or are associated with northeast Oklahoma City.

Table II.1. Description of intended intervention and counterfactual components and target populations

Component	Curriculum and content	Dosage and schedule	Delivery	Target Population
Intervention				
CORE activities required of all participants assigned to the intervention				
TD-OMS	*Newly created adaptation of On My Shoulders (PREP) and other evidence-based interventions. Includes material on couple communication, parenting, social support, and economic stability*	*6 3-hour group workshops (18 hours total) over 6 weeks*	*Workshops offered in person and led by two trained educators (typically male-female team)*	*Low-income dad and co-parent who are raising at least one young child*
Case Management	*Intake meeting for needs assessment; subsequent visits as requested by participants or by educator referral*	*Minimum 1 hour each visit*	*Individual meeting of dad or co-parenting team with coach*	*Low-income dad and co-parent who are raising at least one young child*
Optional INTENSIVE WORKSHOPS				
Relationship intensive	*PREP 8.0, an evidence-based couple communication program. Includes material on understanding partner's perspective; avoiding destructive conflict; and communicating effectively*	*6 3-hour group workshops (18 hours total) over 6 weeks*	*Workshops offered in person and led by two trained educators (typically male-female team)*	*Low-income co-parenting teams who completed TD-OMS*

Component	Curriculum and content	Dosage and schedule	Delivery	Target Population
Parenting intensive	*Newly created parenting curriculum adapted from evidence-based programs*	*6 3-hour group workshops (18 hours total) over 6 weeks*	*Workshops offered in person and led by two trained educators (typically male-female team)*	*Low-income co-parenting teams who completed TD-OMS*
Employment track	*Resume preparation; interview and communication skills; appropriate work attire; financial literacy; job placement assistance*	*10 daily sessions/ 6 hours each (up to 60 hours total) over 2 weeks*	*Multiple staff involved*	*Fathers only*
Counterfactual				
Business as usual: No workshops or supportive services were offered to co-parenting teams in the control group by TRUE Dads				

TD-OMS = TRUE Dads On My Shoulders; PREP = Prevention and Relationship Education Program

5. Education and training of staff

Workshops were delivered by two trained educators, typically a female-male team. Educators were required to have at minimum a bachelor's degree, prior experience in public speaking, and effective presentation skills (see Table II.2). Four Educators out of 10 had advanced degrees in a human services or related areas. They receive additional on-the-job training and quarterly refresher trainings. Educators also completed peer evaluations after each session to support parents' continued growth. Periodically, supervisors and/or local evaluators observed their workshops to provide direct feedback. Educator teams rotated through TD-OMS and the parenting and relationship intensive workshops. At the end of the program's 5th year, 70% of the original educators and 67% of the original coaches were still delivering the TRUE Dads intervention.

Case Managers met with each co-parenting team at the beginning of their participation, and subsequently as needed, to boost participation, assess needs, and refer to other services in the community as needed. They had a bachelor's degree in sociology or psychology and over 13 years of experience in case management, coaching, or equivalent field, and participated in subsequent training each year of the project.

WorkForward coaching staff were trained by and assigned only to that program. The three coaches had a minimum of a bachelor's degree in sociology or psychology and/or over 13 years equivalent experience in case management and/or coaching, and held the following certifications: Civil Mediation, Family/Divorce mediation, Job Coach, Employment Specialist, and Career Consultant. They received multiple trainings that included: Domestic Violence, Motivational interviewing, Understanding Mental Illness, Suicide prevention, Bridges out of Poverty, Diversity and Cultural competency, Psychology of Hope, and Adverse Childhood Experiences.

Table II.2. Staff training and development to support intervention components

Component	TCEducation and initial training of staff	Ongoing training of staff
TD-OMS	Educators have a minimum of a bachelor's degree and initial training of 30 hours over 4 days	Group meetings every two weeks, quarterly trainings (half-day), occasional feedback from staff observer
Relationship intensive	Educators receive relationship intensive training of 12 hours over 2 days	Group meetings every two weeks, one refresher training, occasional feedback from staff observer
Parenting intensive	Educators receive parenting intensive training of 18 hours over 3 days	Group meetings every two weeks, one refresher training, occasional feedback from staff observer
Employment track	Employment specialists and coaches in WorkForward receive training of 20 hours over 3 days	Observation and feedback from a supervisor, feedback from the training team
Case management	A bachelor's degree in sociology or psychology and over 13 years of experience in case management, coaching, or equivalent field	Over the course of 5 years, coaches received more than 100 hours of training per year in a variety of workshops listed in the main text (see section II.A.5)

B. Description of counterfactual condition as intended

Fathers and co-parents in the control condition received no family-based intervention services from TRUE Dads or from other projects operated by IMCI, or Public Strategies, which are located in the same building. Parents in the control group received no information from TRUE Dads staff about family-based community services but were free to use other resources available in the community (for example, Women, Infants, and Children (WIC) or other welfare programs; financial counseling; employment services; individual or family therapy; or parenting classes). No other fatherhood services are available in Oklahoma County or surrounding counties, and fatherhood services focused on co-parenting teams are rare throughout Oklahoma.

C. Research questions about the intervention and counterfactual conditions as implemented

In Table II.3 we list our research questions concerning the implementation of TRUE Dads. Because this program is new, we first examined whether the program was able to attract co-parenting teams and implemented as intended. Additional topics covered include program dosage and attendance, qualifications of educators, quality of the program, participant engagement, and whether similar programming outside of TRUE Dads was available to co-parenting teams assigned to the control condition. Specific measures to answer each of these questions are described later in this report.

Table II.3. Research questions about implementation of TRUE Dads

Implementation element	Research question
Recruitment	• As a new program, can we successfully recruit parenting teams to TRUE Dads? • Will prospective participants accept random assignment?
Fidelity	• Did the educators have adequate educational background to deliver this program? • What steps were taken to modify/improve TRUE Dads prior to the impact study? • Did the educators deliver the curriculum content as intended to the participants?

Implementation element	Research question
Dosage	• How often did the intervention group participate in the intervention on average? • Did attendance meet benchmarks: 60% of the co-parenting teams had at least one member attending 5 of 6 core sessions and 50% of the teams had at least one member (either father or co-parent) attending all 6 sessions?
Quality	• What mandatory training did the educators have? • What steps (for example, supervision or additional training) were taken to improve the quality of service delivery? • Was the program perceived positively by the participants?
Engagement	• How engaged were participants in the intervention?
Context	• What other similar healthy marriage/relationship or responsible fatherhood programming was available to control group members?

III. Study design

This section provides a description of the study design, the sample, and the data collection process for the impact and implementation analyses. The research design and informed consent process was approved by the Committee for the Protection of Human Subjects at the University of California, Berkeley (Protocol ID 2016-08-9047, June 02, 2017).

A. Sample formation and research design

1. Eligibility and recruitment

Recruitment for the study began in the spring of 2017 and ended in June 2019. TRUE Dads staff targeted recruitment toward low-income dads, particularly men of color, who: (a) had a child from birth-12 years old, with special emphasis on ages 0-5; (b) were able to involve a co-parent in the study; (c) were age 18 or older; and (d) resided in or were associated with northeast Oklahoma City, a disadvantaged, predominantly African-American community. Each father who met these criteria was required to bring a co-parent who was directly participating in raising his youngest child. This co-parent could include their spouse, a current or former romantic partner, a family member, or a friend. In all, 1.8% of the co-parents were male, with most of them in non-romantic relationships (uncle, friend, father) with the fathers. Participants were recruited through radio advertising, outreach by intake workers, and referrals from friends. Primary outreach locations included WIC clinics, public libraries, schools, parole and probation offices, and local events.

Many of the participants had more than one child. In order to manage the data from more than 1,000 families, we asked co-parents to focus on their youngest child (the "focal child") when reporting child outcomes in surveys and focus groups. The focal child was not the sole focus of co-parents' participation in the intervention as many co-parents also raised issues and concerns in group meetings about parenting their older children.

Two-thirds of the focal children were biological offspring of both participating co-parents. In 20% of co-parenting teams, the father was the focal child's biological parent but the co-parent was not. In 17% of pairs, the co-parent was the focal child's biological parent but the father was not (child born from co-parent's prior relationship).

2. Consent process

Initial interviews to determine enrollment in the study took place from July 2017 to June 2019. At the end of the intake interview, an IMCI Community Relations and Intake Specialist handed participants the consent form approved by the institutional review board, read the consent form to them, and answered any questions before asking them to agree and sign.

3. Random assignment

Co-parenting teams came together for an intake interview in which the intake worker used a mixture of verbal program descriptions and video presentations. The worker then asked parents

to consent to participate in both the intervention program and the study. Staff then administered three surveys on iPads: ACF's Applicant Characteristics; ACF's Entrance Survey; and the local evaluation's Baseline questionnaire. After intake, consent, and initial data collection, the intake worker randomly assigned co-parenting teams with a random assignment algorithm using a 60/40 (intervention/control) probability. Neither staff intake workers nor participants were aware of the outcome of random assignment during the intake visit. For those assigned to the intervention, the random assignment outcome was conveyed to a coach who contacted participants within 24 hours. For those assigned to the control condition, a staff research specialist notified participants. In order to maintain a flow of participants into the TRUE Dads group, this process was completed on a rolling basis for every 50 co-parenting teams.

B. Data collection

We discuss data collection, first for the implementation analysis and then for the impact evaluation.

1. Data Collection for Implementation analysis

Our implementation research questions from Table II.3 concern recruitment, educator qualifications, treatment fidelity and quality, dosage, participant engagement, and the community context in which the intervention occurred. We rely on six major data sources that yield both quantitative and qualitative insights to answer these questions. The qualitative data were not derived from formal observation instruments but rather from verbal exchanges in regular meetings, conference calls, and meeting notes, and shared with the participating staff.

1. Qualifications of the group educators were obtained from their personnel files. The quality of their work was assessed using impressions during ongoing training, individual and group supervision, and occasional observations of videos of work in the groups.
2. Data on co-parenting team recruitment and participation were obtained from IMCI's management information system (MIS). These records provided data on co-parenting team recruitment, demographic data, surveys completed, attendance at workshops, and contacts with coaches.
3. Co-parenting teams' assessments of the program, their relationships with the educators, and how much they had learned were obtained from participants' survey responses.
4. Qualitative data were gathered from senior staff on an ongoing basis through regular staff meetings and observations of the educator training sessions. These data were in the form of program impressions shared with the evaluators in regular conference calls – every two weeks in the first months, and once monthly after that. Staff providing these data included program administrators and managers, group leaders, recruiting and intake staff.
5. Qualitative data were gathered through observations of TRUE Dads workshop sessions conducted by program staff using detailed notes, followed by direct feedback to educators.
6. Qualitative data were gathered from structured interviews and focus group meetings with participants and staff in the second and fourth year.

Table C.1 in Appendix C describes the data source, timing/frequency of data collection for each implementation question listed in Table II.3 above, and the party responsible for data collection.

2. Data Collection for Impact Analysis

Five sources of participant data were gathered using iPads or collected online and provided to the evaluators at the University of California, Berkeley. Table III.1 summarizes the key features of the data collection. The data were administered and collected in the same way for participants in the intervention and control conditions. Evaluators received only deidentified data with participants' code numbers in order to link files across data sources. The five data sources were:

1. An Applicant Characteristics and Entrance Survey, submitted through Information, Family Outcomes, Reporting and Management (nFORM). nFORM is a cross-site MIS used by grantees of the Office of Family Assistance.
2. A 600+ item baseline assessment completed by each member of the co-parenting team. This assessment was administered in Qualtrics, a survey software program.
3. Supplementary data and participant demographics gathered from the MIS.
4. Exit surveys collected after the final TD-OMS workshop session and again after the final intensive workshop. These surveys were collected in nFORM and in the MIS, respectively. We did not use these data in our impact analysis because they were collected only for the intervention group and exhibited a low response rate.
5. A follow-up survey administered one-year after baseline (approximately 8 months after most participants had completed the intervention). These surveys were administered using Qualtrics and in person at IMCI on an iPad or, if requested, at another location of the participants' choosing. Starting in March 2020, the survey was administered online due to the global coronavirus pandemic. For program participants, the survey also asked about their program experiences.

Table III.1. Key features of the data collection for the impact analyses

	Data source	Timing of data collection	Mode of data collection	Party responsible for data collection	Start and end date of data collection
Intervention	Intervention fathers and co-parent	Enrollment (baseline)	• In-person via iPad: • nFORM Applicant Characteristics, • nFORM Entrance Survey, • Qualtrics baseline survey	Program staff	June 12, 2017 through June 30, 2019
	Intervention fathers and co-parent	Follow-up (1 year after baseline; approximately 8 months after program completion).	• In-person via iPad or online: • Qualtrics follow-up survey	Program staff	June 12, 2018 through June 30, 2020

	Data source	Timing of data collection	Mode of data collection	Party responsible for data collection	Start and end date of data collection
Counterfactual	Comparison father and co-parent	Enrollment (baseline)	• In-person via iPad: • nFORM Applicant characteristics, • nFORM Entrance Survey, • Qualtrics baseline survey	Program staff	June 12, 2017 through June 30, 2019
	Comparison father and co-parent	Follow-up (1 year after baseline; approximately 8 months after program completion).	• In-person via iPad or online: • Qualtrics follow-up survey	Program staff	June 12, 2018 through June 30, 2020

nFORM = Information, Family Outcomes, Reporting and Management

Dads and co-parents in the intervention and control groups each received $30 for responding to the Applicant Characteristics and Entrance Survey, $40 for the exit survey (intervention group only), and $50 for the follow-up survey. Six months after baseline, all participants were reminded about the upcoming follow-up survey when they were contacted by TRUE Dads research staff.

IV. Analysis methods

In this section we describe the construction of the sample used for analysis, the outcome measures, and the baseline equivalence of participants randomly assigned to the treatment and control conditions.

A. Analytic sample

TRUE Dads randomly assigned 1,042 co-parenting teams, with 621 (59.6%) invited to participate in the TRUE Dads program and 421 (40.4%) assigned to the control group who did not receive services from TRUE Dads during the study period (see CONSORT diagram in Appendix B).

The overall response rate for the 891 co-parenting teams in which at least one partner completed the 1-year follow-up survey was 85.5%. After accounting for co-parenting team attrition, the individual response rate, measured from the sample of non-attriting couples, was 90.1%. While there was little variation from the overall response rate for measures of depression, destructive communication, and control violence, 83 fewer co-parenting teams had baseline and follow-up responses to the survey questions regarding harsh parenting and child behavior because they entered the study before the birth of their child; the response rate for these measures was 77.5%.

We then turned to an analysis of possible differential response rates in the intervention and control samples. In 531 of the 621 co-parenting teams assigned to the intervention, and 360 of the 421 teams assigned to control, the overall co-parenting team attrition in the intervention and control conditions was identical (85.5%). Looking at each of the 6 outcome measures separately, we found that for couple outcome measures of depression, destructive communication, and control violence, there was no difference or fractional difference (.05%) between the retention rates for intervention and control samples. Two outcome measures, harsh parenting and child behavior problems showed a differential retention differential of about 3% in favor of TRUE Dads participants. The one outcome variable measured at the individual level – fathers' employment - showed slightly higher response retention rate (87.4%) in non-attriting clusters than the best retention rate of couple data (85.5%) and a higher experimental vs control differential attrition rate, just over 5% *(Chi-square* = 4.90 p = .027). This difference may be an intervention effect, in that participation in TRUE Dads encouraged slightly more of the men to report their employment. Despite this rather small difference, the overall attrition was low (and conversely, retention rate was high) and there was little difference in the rate of attrition between intervention and control participants. We can conclude this study has low attrition according to What Works Clearinghouse (2020) standards.

The TRUE Dads intervention chose to approach the goal of enhancing father's positive involvement in family life using couples' groups as an intervention modality. As evaluators, we chose to analyze the impact of the intervention using couple rather than individual data except for fathers' employment. We provide a detailed discussion of the rationale for these choices in Appendix E.

The study analyzed all co-parenting teams according to their randomly assigned condition, regardless of whether they participated actively in TRUE Dads workshops—an intent-to-treat design.

Table IV.1. Cluster (couple) and individual sample sizes by intervention status

Number of:	Intervention sample size	Comparison sample size	Total sample size	Total response rate	Intervention response rate	Comparison response rate
Clusters: At beginning of study	621	421	1,042	n.a.	n.a.	n.a.
Clusters: Contributed at least one individual at baseline	621	421	1,042	100%	100%	100%
Clusters: Contributed at least one individual at 1-year follow-up)	531	360	891	85.5%	85.5%	85.5%
Clusters: Contributed to the impact analysis for outcome: depression	531	360	891	85.5%	85.5%	85.5%
Clusters: Contributed to the impact analysis for outcome: destructive communication	525	355	880	84.5%	84.5%	84.3%
Clusters: Contributed to the impact analysis for outcome: coercive violence	525	355	880	84.5%	84.5%	84.3%
Clusters: Contributed to the impact analysis for outcome: harsh parenting	490	318	808	77.5%	78.9%	75.5%
Clusters: Contributed to the impact analysis for outcome: child behavior problems	490	318	808	77.5%	78.9%	75.9%
Individuals in non-attriting clusters						
Individual: At time that clusters were assigned to condition	1,062	720	1,782	n.a.	n.a.	n.a.
Individual: Who consented	1,062	720	1,782	100%	100%	100%
Individual: Contributed a baseline survey	1,062	720	1,782	100%	100%	100%
Individual: Contributed to 1-year follow-up	967	639	1,606	90.1%	91.1%	88.8%
Individual: Contributed to the impact analysis for outcome: depression	967	639	1,606	90.1%	91.15%	88.8%
Individual: Contributed to the impact analysis for outcome: destructive communication	955	633	1,588	89.1%	89.9%	87.9%

Number of:	Intervention sample size	Comparison sample size	Total sample size	Total response rate	Intervention response rate	Comparison response rate
Individual: Contributed to the impact analysis for outcome: coercive violence	958	634	1,592	89.3%	90.2%	88.1%
Individual: Contributed to the impact analysis for outcome: harsh parenting	935	614	1,549	86.9%	88.0%	85.3%
Individual: Contributed to the impact analysis for outcome: child behavior problems	936	611	1,547	86.8%	88.1%	84.9%
Fathers in non-attriting clusters						
Fathers: contributed to 1-year follow-up	531	360	891	100%	100%	100%
Individual: Contributed to the impact analysis for outcome: father's employment	475	304	779	87.43%	89.5%	84.4%

n.a. = not applicable.

B. Outcome measures

Consistent with our multi-domain model of the risks and protective factors affecting individual, couple, and family adaptation, we selected measures of (a) individual functioning of father and co-parent, (b) co-parenting team relationship quality, (c) parent-child relationship quality,

(d) child behavior problems, and (e) fathers' economic self-sufficiency. All of our multi-item scales have adequate reliability (Cronbach's alpha greater than .70). Fathers' economic self-sufficiency was analyzed at the individual level. The remaining 5 measures were analyzed using the couple-level average.

1. Primary outcome measures

Table IV.2. Outcome measures used for primary impact analyses research questions

Measure	Description of the outcome measure	Source of the measure
Symptoms of Depression	Center for Epidemiological Studies in Depression scale (CES-D); sum of 12 items reported using 0 (rarely) to 3 (most or all of the time); scale range 0 to 36. Sample item: "I felt sad." Cronbach's alpha: .91 (father), .93 (co-parent).	Brief version of CES-D (Radloff, 1977)
Co-parents' relationship conflict	Destructive communication scale; mean of 9 items reported using 0 (never) to 3 (often); scale range 0 to 3. Sample item: "Little arguments turn into ugly fights." Cronbach's alpha: .95 (father), .96 (co-parent).	PREP
Co-parents' relationship domestic violence	Domestic Violence: Coercive Control scale; mean of 8 yes (1) or no (0) items; scale range 0 to1. Sample item: "Partner threatened to hurt you or your child." Cronbach's alpha: .80 (father), .80 (co-parent).	Investigator-created measure adapted from Conflict Tactics Scale (Straus)

Measure	Description of the outcome measure	Source of the measure
Parenting quality: Harsh parenting	Harsh Parenting scale; average of 6 items reported using 0 (not in the past month) to 4 (every day or almost every day); scale range 0 to 4. Sample item: "How often have you yelled at target child?" Cronbach's alpha: .89 (father), .86 (co-parent)	Investigator-created measure adapted from Alabama Parenting Scale and Child Trends survey items
Child behavior problems (focal child)	Child Adaptive Behavior Inventory describing focal child behavior problems: angry, hyperactive, sad/anxious, and shy/withdrawn behavior; average of 27 items reported using 1 (not at all like this) to 4 (very much like this); scale range 1 to 4. Sample items: "Gets into fights", "has trouble concentrating." Focal child is youngest child of co-parents. Cronbach's alpha: .78 (father), .78 (co-parent)	Child Adaptive Behavior Inventory (Cowan, Cowan, & Heming, 2005)
Father's Employment	Employment single item (value range: 1= Not currently employed; 2= Temporary, occasional, or seasonal employment, or odd jobs for pay; 3 = Employed, but number of hours changes from week to week; 4=Part-time employment (usually work 1 – 34 hours a week); 5=Full-time employment (usually work 35 or more hours a week).	Investigator-created measure

Notes: PREP = Prevention and Relationship Education Program. Outcomes were measured 12 months after baseline.

1. Additional outcome measures

As noted in our description of the curriculum, the TRUE Dads intervention addressed six risk and protective domains of family life (See Appendix A for our theory of change). Our additional analysis traces the pathways through which the intervention produces positive changes in each of the domains: life stressors, personal distress, co-parents' relationship quality, parenting quality, child behavior problems, and fathers' economic self-sufficiency.

Our secondary analysis was conducted with Structural Equation Modeling, using latent variables rather than single scales to provide measures of each construct. For example, instead of assuming that a measure of depressive symptoms adequately represents the construct "personal distress," we combined scales measuring parents' depressive symptoms, anxiety, and anger. The advantage of latent variables over single measures is that they create estimates of complex constructs, somewhat like factors, with each variable weighted according to its contribution to a statistically estimated construct. That is, the measures are not simply added together to produce an arithmetic sum or averaged to produce a mean score. Consistent with our couple-level approach to analysis, each construct, except for fathers' economic self-sufficiency, contained both partners' responses in each co-parenting team. The 6 latent variable constructs and the 22 measures included in our Structural Equation Model are described in Table IV.3 (also see Figure 1).

Table IV.3. Outcome measures used for secondary impact analysis research questions

Outcome measure	Description of outcome measure	Source
Life stressors (3 scales from each partner)	*Life Stress Events*. 14 items divided into three scales (Value range 1-4; never, once or twice, several times) e.g.," During the past month I have experienced stress resulting from difficulties at work/housing difficulties, a physical or mental illness in my family, difficulties with friends." Cronbach's alpha =.77 (father), .76 (mother)	Adaptation of Holmes and Rahe Life Stress Events Scale
Personal Distress (3 scales from each partner)	*Center for Epidemiological Studies in Depression scale (CES-D) 12 items (value range 1 to 4; rarely* to most or all the time). e.g., "I felt sad". Scale score range 12-48. Cronbach's alpha =.91 (father), .93 (mother)	Brief version of the widely used CES-D (Radloff, 1977)
	Anxiety 6 items (value range 0 to 3; "not at all" to "yes very often") "In the past 7 days I have been anxious or worried for no good reason". Cronbach's alpha =.82 (father), .86 (mother).	K. Edin
	Anger. 5 items. (Value range 1-5; not at all to very much.) "I felt angry."	Child Trends
Co-parents' relationship quality (4 scales from each partner)	*Destructive communication scale*; mean of 9 items rating 0 (never) to 3 (often); scale range 0 to 3. e.g., "Little arguments turn into ugly fights." Cronbach's alpha = .95 (father), .96 (mother).	*PREP*
	Constructive communication scale: mean of 8 items rating 0 (never) to 3 (often); scale range 0 to 3. e.g., "Even when arguing we can keep a sense of humor" Cronbach's alpha = .95 (father), .96 (mother).	*PREP*
	Problem solving style: mean of 7 items rating 1(never) to 7 (always). Please indicate how often each statement is true in your relationship. e.g., "When discussing problems, we work together as a team until we have a solution.) Cronbach's alpha = .91 (father), .92 (mother).	*PREP*
	Happiness with the relationship. 1 item. "Taking all things together, on a scale from 0 to 10, where 0 is not at all happy and 10 is completely happy, how happy would you say your co-parenting relationship with your partner is? "	Common item in marital satisfaction questionnaires
Parenting quality (3 scales from each partner)	*Parenting Stress Index* 4 items. (Value range 0-4. Strongly disagree to strongly agree). "Child turned out to be more of a problem than I expected" Cronbach's alpha =.78 (father .79), .82 (mother).	Abidin PSI
	Harsh Parenting 6 items. Value range 0-4. From "not in the past month" to "every day or almost every day." e.g., "How often have you yelled at target child?" "How often have you hit target child?" Cronbach's alpha =.84 (father), .86 (mother).	Alabama Parenting Questionnaire
	*Father involvement 3 scales Value range of each 1-9) 1=he does none (e.g., feeding the child, taking the child to the doctor) 9=co-parent does it all. Value range of combined scale 1-*9. Cronbach's alpha =.88 (father), .82 (mother).	Who Does What? (Cowan & Cowan)
Child behavior problems (4 scales from each partner)	*Child behavior problems. The 27 items were divided into 4 child behavior problem scales: angry, hyperactive, sad/anxious, and shy/withdrawn behavior (value range 1 to 4, not at all like this to very much like this) e.g., "gets into fights", "has trouble concentrating", "cries a lot", "is shy or bashful with other children."* Average Cronbach's alpha across scales =.91 (father), .91 (mother).	*Child Adaptive Behavior Inventory (Cowan, Cowan, & Heming 2005)*

Father's economic self-sufficiency *(5 scales from fathers only)*	*Father's employment 1 item (value range 1-5 from "not currently employed" to "full time employment.")*	Investigator-created measure
	Father's income. Single item. What was your total individual income from all jobs in the last 12 months before taxes? Scale score from 0 to 17 in $5,000 increments.	Investigator-created measure
	Income. Single item. "What was your total individual income from all jobs in the last 12 months before taxes?" Scale score from 0 to 17 in $5,000 increments.	Investigator-created measure
	`Employment Hope Scale 6 items (Value range from 1-7, disagree to agree). "I feel positive about how I will do in my future job situation" (Cronbach's alpha =.94, .95)	*From Hope Scale (Hong)*
	Employment confidence 4 items. (Value range from 1-4. Strongly disagree to strongly agree) "I feel confident about my interviewing skills" (Cronbach's alpha =.75, .79)	*From Hope Scale (Hong)*

Notes: ` PREP = Prevention and Relationship Education Program. Outcomes were measured 12 months after baseline. Except for fathers' economic self-sufficiency, each measure was administered to both co-parents, and the latent variable contains both co-parents' responses.

C. Baseline equivalence and sample characteristics

We have named the analytic sample of 891 co-parenting teams in which at least one partner provided 1-year follow-up data on the CES-D depression scale, the Benchmark sample (with the highest response rate). For this Benchmark sample, we tested whether there were statistically significant differences between intervention and control groups at baseline in key demographics and primary outcome measures (See Table IV.4a). We note that in this sample there were 11 fewer co-parenting teams who completed the destructive communication and control violence measures at follow-up, and 83 fewer teams who completed questionnaire measures of harsh parenting and child behavior problems. This latter result was attributable to the fact that these parents filled out baseline measures at study entrance when the child had not yet been born.

Baseline differences and *p*-values for dichotomous data were calculated with a Chi-square test, and effect size estimates were calculated using the Cox index. Baseline differences and *p*-values for continuous data were calculated with *t*-tests, and effect sizes estimated using Hedges-*g*. Details of the rationale for our baseline equivalence analysis can be found in Appendix F.

For the Benchmark sample, we examined baseline differences between TRUE Dads participants and control condition participants for 14 demographic variables, and 5 outcome measures. *There were no statistically significant baseline differences between intervention and control participants for any of these measures (see Table IV.4a)*. Two demographic variables had nonsignificant effect sizes over .10. The TRUE Dads intervention group had a slightly lower proportion of Hispanic participants (Cox $d = .19$) and co-parents with a high school diploma or above (Cox $d = .14$).

A table describing baseline equivalence for the analytic sample of fathers who responded to questions about employment is presented in Table IV.4b. Again, there were no statistically

significant baseline differences in any of the measures. In this Table only one effect size was over .10 – the proportion of co-parents with high school degrees or above (Cox *d* =.16).

We examined baseline differences separately for analytic data sets based on the slightly smaller samples associated with destructive communication and control violence (n=880; See Appendix Table F.1) and for harsh parenting and child behavior problems (n=808; See Appendix Table F.2). Each of these had one of 19 statistically significant findings: In Table F.1, there were significantly more non-intimate partners in the experimental group at baseline, while in Table F.2 the intervention group had significantly fewer Hispanic participants.

We draw four conclusions from this pattern of baseline difference analyses. First, there were no statistically significant differences between intervention and control groups or effect sizes above .10 for the 6 primary outcome measures of this study. Second the two statistically significant baseline non-equivalences were found in distributions in which the proportions of Hispanic participants and the proportion of non-intimate partners was small; the Chi-square test exaggerates the effects when the variable is distributed unequally. Third, over the four Tables we documented 76 statistical tests of baseline differences and found only two that were statistically significant – a result that was less than that expected by chance. Fourth, as a result of these analyses we included the three variables (Hispanic, non-intimate partner, and co-parent with a high school diploma or above) that showed effect sizes higher than .10 in sensitivity analyses in order to test the generalizability of the primary impact of the TRUE Dads intervention.

Table IV.4a. Key baseline measures and baseline equivalence across study groups for 891 co-parenting teams

Baseline measure	Intervention mean (standard deviation)	Comparison mean (standard deviation)	Intervention versus comparison mean difference (p-value of difference)	Effect size
Fathers (%)	50	50	0	0
Fathers' Race/ethnicity (%)				
Hispanic	9.0	11.1	-2.1 (*p*=.089)	.19
Native American	13.9	14.1	-.2 (*p*=.631)	.05
Non-Hispanic White	27.9	26.8	-1.1 (*p*=.863)	.01
Non-Hispanic Black	43.3	41.4	1.9 (*p*=.922)	.01
Married (%)	34.0	30.7	-3.3(*p*=.306)	.09
Living together (%)	84.7	84.2	.5 (*p*=.837)	.03
Father is unemployed (%)	38.9	39.1	-.2 (*p*=.539)	.04
Father has high school diploma or above (%)	63.4	61.2	2.2 (*p*=.516)	.06
Co-parent has high school diploma or above (%)	70.2	74.8	-4.6 (*p*=.150)	.14
Father's annual income ($)	10,850 (10,986)	10,150 (10,756)	700 (p=.517)	.06
Co-parent's annual income ($)	8,340 (8,526)	7,600, (8263)	740 (*p*=.382)	.09
Father's age	31.76 (7.54)	31.67 (7.67)	-.09 (*p*=.852)	.01

Baseline measure	Intervention mean (standard deviation)	Comparison mean (standard deviation)	Intervention versus comparison mean difference (p-value of difference)	Effect size
Co-parent's age	30.92 (8.58)	30.81 (8.21)	.11 (*p*=.837)	.01
Youngest Child's age	3.02 (3.15)	2.77 (3.06)	.25 (*p*=.252)	.08
Baseline scores for the primary outcome variables for co-parenting teams in which either partner completed 1-year follow-up				
Depression[1]	10.07 (5.86)	9.71 (6.19)	. 36 (*p*=.450)	.06
Destructive communication[2]	1.36 (.79)	1.37 (.77)	-.01 (*p*=.497)	.01
Coercive violence[2]	.26 (.23)	.27 (.23)	-.01 (*p*=.740)	.04
Harsh parenting[3]	.73 (.71)	.67 (.69)	.06 (*p*=.238)	.09
Child behavior problems[3]	1.82 (.39)	1.80 (.39)	.02 (*p*=.508)	.05

Notes: *p*-values (two-tailed) are included in parentheses. Effect sizes are calculated by Cox index (dichotomous data) and Hedges-*g* (continuous data).

[1] Depression: Intervention n=531; Control n=360

[2] Destructive communication and Coercive violence: Intervention n=525; Control n=355

[3] Harsh parenting and Child behavior problems: Intervention n=490; Control n=318

Table IV.4b. Key baseline measures and baseline equivalence across study groups for 779 fathers

Baseline measure	Intervention mean (standard deviation)	Comparison mean (standard deviation)	Intervention versus comparison mean difference (p-value of difference)	Effect size
Fathers' Race/ethnicity (%)				
Hispanic	9.1	10.0	-.9 (*p*=.441)	.10
Native American	14.0	14.2	-.2 (*p*=.660)	.05
Non-Hispanic White	28.7	28.0	.7 (*p*=.834)	.02
Non-Hispanic Black	42.3	41.7	.6 (*p*=.614)	.05
Married (%)	35.4	32.5	2.9 (*p*=.402)	.08
Living together (%)	85.2	83.9	1.3 (*p*=.668)	.06
Father is unemployed (%)	37.5	37.2	.3 (*p*=.929)	.01
Father has high school diploma or above (%)	65.0	62.1	2.9 (p=.428)	.08
Co-parent has high school diploma or above (%)	70.1	75.3	-5.2 (*p*=.145)	.16
Father's annual income ($)	10,950 (10,744)	10,450 (11,058)	500 (*p*=.687)	.03
Co-parent's annual income ($)	8,400 (8650)	8,150 (8583)	350 (*p*=.816)	.02
Father's age	31.88 (7.72)	31.99 (7.78)	-.11 (*p*=.847)	.01
Co-parent's age	30.91 (8.58)	31.12 (8.50)	-.21 (*p*=.739)	.02
Youngest Child's age	2.99 (3.18)	2.87 (3.15)	.12 (*p*=.618)	.04
Baseline scores for the primary outcome variable				
Father's employment[1]	2.88 (1.79)	2.97 (1.79)	*-.9 (p=.486)*	.05

Notes: *p*-values (two-tailed) are included in parentheses. Effect sizes are calculated by Cox index (dichotomous data) and Hedges-*g* (continuous data).

[1] Father's employment: Intervention n=475; Control n=304; 1= Not currently employed; 2= Temporary, occasional, or seasonal employment, or odd jobs for pay; 3 = Employed, but number of hours changes from week to week; 4=Part-time employment (usually work 1 – 34 hours a week); 5=Full-time employment (usually work 35 or more hours a week).

Not shown in Table IV.4a and IV.4b are the baseline equivalence analyses for the 6 latent variables from Table IV.3. None of the baseline differences between control and intervention conditions was statistically significant for these 6 variables, and effect sizes ranged between .00 and .03.

In sum, we conclude that the randomization procedures worked well and that the intervention and control samples were relatively equivalent at baseline. Nevertheless, we re-examine this hypothesis using sensitivity analyses in section V.C below.

V. Findings and Estimation Approach

A. Implementation evaluation

Key Findings:

TRUE Dads was successfully implemented with fidelity to the intentions of the program designers. 1,042 co-parenting teams were recruited. More than 80% of the teams attended at least one core workshop session, and almost half (49%) attended all 6. More than 90% of fathers and co-parents rated their experience in the program positively.

Given that TRUE Dads was a new, innovative co-parenting team approach to responsible fatherhood interventions, our main implementation questions concerned whether it was possible to mount the planned 16-week program, attract co-parenting teams to participate, recruit and train a skilled staff of educators and coaches, and deliver a high-quality program with fidelity that would engage participants who valued their experience with the educators and with other group participants. Details concerning the six implementation elements presented in Table II.3 above are addressed in this section. Based on our assessment of data on recruitment, fidelity, dosage, quality, engagement, and context we conclude that the implementation of TRUE Dads was successful. For further information on implementation measures, see Table B.1 in Appendix B.

1. Recruitment

IMCI staff were successful at recruiting fathers who were able to convince the co-parents of their youngest child to join them in participating in the research project. Furthermore, almost all of the co-parenting teams who came for their intake interview agreed to accept random assignment to TRUE Dads or control conditions. IMCI staff recruited 1,042 eligible co-parenting teams during the two-year study period, exceeding the recruitment target of 1,000 co-parenting teams. Teams were assigned as planned, with 60% randomly assigned to the intervention and 40% to the control group.

2. Fidelity

The program was implemented with fidelity. Table II.2 shows that we were able to recruit a well-educated cadre of educators and coaches who were given training from the program developers at the beginning of the project, regular supervision, and systematic refresher training over the course of the study. Educators also used a fidelity tool during each group meeting to stay on track with session goals and outputs. Supervisor reports and ongoing observation of the groups with direct feedback to the educators also served to keep the educators faithful to the curriculum as intended. Regular meetings with educators and coaches, senior staff meetings, and conference calls between IMCI staff and evaluators resulted in small changes in implementation during the pilot phase. For example, in response to slow intake, IMCI added radio advertising (paid) and participant interviews (free) on a station with high listener participation in the communities. Data

collection procedures were streamlined. Meetings during the pilot phase with educators clarified how modules that were confusing to them were supposed to be administered. Adjustments were made in the amount of material covered each week to make certain that conveying concepts fully was given priority over strict coverage of the materials in the curriculum manual.

3. Dosage

During the planning phase of the project, we set program goals for attendance: 60% attendance at 5 of 6 core sessions and 50% attendance at all 6 core sessions. For the 6 core workshop sessions, 81% of the teams attended one or more; 67% attended at least half; 59% attended 5 of 6 (the first benchmark); and 49% attended all 6 (the second benchmark). Of the 6 optional intensive workshops on couple relationships or parenting, 27% of teams attended three or more, and 17% attended all 6.

As frequently happens in interventions offered to highly stressed low-income families, 19% of those who agreed to the random assignment to TRUE Dads did not attend a workshop. If we focus only on those who came to the first of the 6 core workshops, 75% attended at least half of them, and 58% attended all 6. Thus, most co-parenting teams that engaged with the program initially experienced a substantial portion of the core program.

4. Quality

Participants' ratings about the quality of their program experience were summarized from two data sources: (1) the follow-up survey administered at 12 months, and (2) their responses in two focus groups of those who had completed the program.

Participants' survey responses. Participants generally rated their experiences with the TRUE Dads program educators positively in the follow-up survey. A 7-item survey scale was constructed to describe participants' working relationship with program staff, with each survey item having a 1-7 scale range (from 1 "strongly disagree" to 7 "strongly agree" with a number of positive statements). Average ratings on this scale were 6.4 for father and 6.3 for co-parent. Sample items on this scale included "The TRUE Dads team worked with us toward goals that make sense to me" and "I felt respected by the TRUE Dads educators." More than 90% of participants rated program quality positively (5 or above out of 7) on all items and 40% rated all seven survey items at 7 out of 7.

Participants' focus group responses. Sociologist Sara Halpern-Meekin and her colleague led focus group interviews in the second and fourth years of the program with two dozen participants who had completed the TRUE Dads program. Their feedback was uniformly positive about their experience in the TRUE Dads program, although several participants wanted more focus on family communication. Information from these focus groups helped to improve some aspects of the program. For example, a few parents were critical of the childcare provisions for their children during the workshops, which led to a change in the childcare providers and the arrangements for childcare.

Two responses reflected a number of the fathers' responses:

> *"I didn't know how to talk about my feelings. I was never taught that. I was taught 'You're the man. You just shut up and deal with it.' And that was like a big rift in our relationship because she would swear I didn't care. So, I've gotten better at not holding it in as much. Now I can come and be like "Baby, my feelings is kind of hurt.'"*

> *"It was just so cool to be together in a room with people who are going through the exact same thing you're going through and trying to get to the same goals that you're getting to."*

5. Engagement

Participants generally rated their experiences with the TRUE Dads program workshops positively in the follow-up survey. Using a 7-item survey scale, fathers' average rating was 6.6; co-parent 6.3. Some sample items on this scale included "I learned new things by being in the group" and "I felt like others in the group cared about me." About 88% of fathers and 91% of co-parents gave positive ratings on each item, and a majority of fathers and co-parents (51% and 61%, respectively) gave the highest rating on all 7 items.

6. Context.

As stated above, during the period of this grant there were no other family-focused fatherhood programs in the Oklahoma or surrounding counties and almost none with a co-parenting focus in the country.

B. Primary impact evaluation

Key Findings:

Participants randomly assigned to the control condition were compared with participants randomly assigned to the TRUE Dads program at 1-year follow-up. Compared to control group members, TRUE Dads fathers and co-parents reported fewer symptoms of depression, less destructive communication, and less controlling violence toward one another. TRUE Dads fathers showed an increase in their level of employment. Two measures did not show direct intervention effects – harsh parenting and children's behavior problems

This section describes the impact of TRUE Dads on our 6 primary outcome measures (See Table V.1). Intervention effects for each of these measures were assessed using Analysis of Covariance (SPSS, version 25). The mean score for the co-parenting team on each outcome measure was used to analyze the five couple level outcomes. For father's employment, the outcome was analyzed considering only the responses of fathers, not the co-parenting team. A probability level of <.05, using a two-tailed test, was the criterion for accepting the hypothesis that the intervention produced a statistically significant impact on TRUE Dads participants. Equations for estimating primary intervention effects are presented in Section E of the Appendix.

Our analyses controlled for the baseline measure of the outcome variables in order to improve the precision of the analytic models. However, because we did not observe statistically significant baseline differences in outcome measures between the TRUE Dads and control group (see Table IV.4a and IV.4b and Appendix F.1 and F.2), we did not include other covariates in our models.

Compared to those assigned to the control condition, one year after enrollment TRUE Dads fathers and co-parents reported:

- Fewer symptoms of depression (F=7.51; p=.006)
- Lower levels of destructive communication (for example, verbal attacks, criticism, or name-calling; F=4.68; p=.031)
- Lower levels of coercive domestic violence (for example, physical arguments, threats of violence to partner or child, or attempts to control partner; F=5.02; p=.025)
- Positive effects on father's level of employment (F=3.78; p=.048).[1]

We found no statistically significant impacts of TRUE Dads on two measures:

- harsh parenting (angry, yelling, screaming, threatening, spanking; F=0.39; p=.534)
- children's behavior problems (aggressive, hyperactive, shy/withdrawn, and sad/anxious behavior; F=0.00; p=.887).

Table V.1. Post-intervention estimated effects on primary outcome measures at 1-year

Outcome measure	TRUE Dads sample size	Control sample size	Intervention mean (standard deviation)	Comparison mean (standard deviation)	Intervention vs comparison mean difference (p-value of difference)	Effect Size
Depression	531	360	8.67 (6.60)	9.73 (7.28)	-1.06 *p*=.006)	.18
Destructive communication	525	355	1.04 (.83)	1.16 (.87)	-.12 (*p*=.031)	.15
Coercive violence	525	355	.18 (.21)	.21 (.25)	-.03 (*p*=.021)	.13
Harsh parenting	490	318	.83 (.76)	.77 (.72)	.06 (*p*=.534)	.08
Child behavior problems	490	318	1.78 (.39)	1.77 (.38)	.01 (*p*=.887)	.03
Father's employment	475	304	3.39 (1.74)	3.16 (1.76)	.23 (*p=.048)*	.18

Notes: See Table IV.2 for a detailed description of each measure.

[1] There was an increase from baseline to follow-up of 14% in TRUE Dads fathers employed at least part time versus an increase of 6% in control fathers employed at least part time.

C. Sensitivity analyses

We conducted sensitivity tests to determine whether the outcome findings were robust across methods of analysis. First, we examined whether the results varied by how we defined the analytic sample. Second, we examined where the impact results were sensitive to the covariates included.

1. Sensitivity to the definition of the analytic sample

In the results above, the Benchmark analytic sample was derived from the mean scores of the parenting teams in which at least one partner completed the one-year follow-up survey. It is also possible to define the analytic sample as including only those co-parenting teams in which *both* partners completed the one-year follow-up survey.

Table V.2 describes the retention rates of the benchmark sample and of the sample in which both partners completed the 1-year follow-up. Two conclusions are clear. First, there is an approximately 20% drop (depending on the measure) in the sample size if we analyze data only from couples in which both completed the 1-year follow-up. This sample size reduction would reduce the power of the analysis to detect intervention effects. Second, there is nearly equal participation in follow-up for intervention and control condition co-parenting teams in which either partner completed the one-year follow-up, but a 4 percentage-point difference in favor of the intervention condition when we look at both partners completing the follow-up. In other words, including only those co-parenting teams in the analysis in which both partners completed the follow-up would have potentially biased the interpretation of the intervention's effects.

Table V.2. Percent of co-parenting teams in which either partner completed one-year follow up (benchmark analytic sample) or both partners completed one-year follow-up for intervention and control conditions

	Either partner completed one-year follow-up (benchmark)		Both partners completed one-year follow-up	
Outcome measure	Intervention (%)	Control (%)	Intervention (%)	*Control (%)*
Depression	86	86	69	65
Destructive communication	85	84	68	64
Coercive violence	85	84	69	65
Harsh parenting	79	76	60	56
Child behavior problems	79	76	61	57

Source: First follow-up surveys administered 12 months after the program.

We conducted impact analyses of our primary outcome variables with the reduced sample in which both partners completed the follow-up. In contrast with the sample in which either partner completed the follow-up (4 of 6 statistically significant intervention effects), only one outcome variable showed a statistically significant intervention effect (See middle columns of Table V.3.) However, the results showed a similar order of the impacts: the effect sizes of the Benchmark sample and the sample in which both partners completed the follow-up were highly correlated

(r=.84). We conclude from this analysis that the difference between the two samples is most likely attributable to the difference in sample size (statistical power).

Table V.3. Differences in means between intervention and comparison groups estimated using alternative methods

Outcome	Benchmark approach At least one partner completes follow-up Baseline as covariate		*Both partners complete follow-up* *Baseline as covariate*		*Benchmark Sensitivity approach* At least one partner completes follow-up Baseline as covariate + 5 additional covariates	
	Difference *p*-value	Effect size	Difference *p*-value	Effect size	Difference *p*-value	Effect size
Couple-level variables						
Depression	-1.06 (*p*=.006)**	.18	-1.17 (*p*=.01)*	.22	-.11 (*p*=.064)+	.11
Destructive communication	- .12 (*p*=.031)*	.15	-.06 (*p*=.204)	.09	-.12 (*p*=.039)*	.14
Coercive violence	.03 (*p*=.021)*	.13	.02 (*p*=.333)	.09	.03 (*p*=.045)*	.14
Harsh parenting	.06 (*p*=.534)	.08	.01 (*p*=.840)	.02	.05 (*p*=.490)	.07
Child behavior problems	.01 (*p*=.887)	.03	.01 (p=.940)	.03	.01 (*p*=.942)	.03
Individual-level variable						
Father's employment	.23 (*p*=.048)*	.18	n/a	n/a	.18 (*p*=.084)+	.11

**/*/+ Differences are statistically significant at the .01/.05/.10 levels, respectively.

2. Sensitivity to inclusion of covariates

For each of the five couple-level primary outcome measures, we repeated the Analyses of Covariance, adding the 5 baseline covariates identified in Table V.4. The first two of these covariates—Hispanic origin and non-intimate partner status -- showed statistically significant differences between intervention and control conditions in one baseline equivalence table (F.1 or F.2). The third co-variate, co-parent with high school diploma or above, consistently showed an effect size greater than .10. The remaining two – whether the couple was married, and father's annual income -- did not show statistically significant baseline differences but were included because of the specific interest of TRUE Dads in whether the marital status of the participants or the income level of the fathers had any effect on the impact of the intervention.

Table V.4. Baseline covariates included in sensitivity analyses

Measures	
Hispanic	Hispanic or not
Non-intimate partner status	Co-parent is a relative, friend of the father, etc.
Co-parent education	High school diploma or above
Marital status	Married or not
Father's income	Annual income

The results of the sensitivity analyses are shown in the right-hand columns of Table V.3. Compared with the Benchmark analysis, which included only the baseline as a covariate, the addition of the 5 covariates had little effect on destructive communication, coercive violence, harsh parenting, or child behavior problems. In addition, the impact of TRUE dads on depression (p=.06) and fathers' employment (p = .08) was now marginally significant when co-variates were added. The correlation between the effect sizes obtained in the Benchmark analysis (Tables V.3 without these 5 covariates) and the effect sizes in the sensitivity analyses was .81. For these outcome variables, then, we can conclude that the impact of TRUE Dads generalizes across several relationship structures of the co-parenting teams as well as the education level of the co-parent and fathers' income.

D. Additional analyses: What does the pattern of outcome changes tell us about the dynamic effects of TRUE Dads on the family and on the fathers' economic self-sufficiency?

Key Findings:

Using a structural equation model, we tested hypotheses about TRUE Dads intervention effects using a latent variable approach. Participation in the intervention was connected to a reduction in personal distress of both fathers and co-parents, and this reduction in parents' depression, anxiety, and anger led to improvements in the quality of the co-parenting relationship. We did not find direct paths linking the intervention with child behavior outcomes but identified indirect effects from parents' intervention participation to parenting quality, which was associated with a reduction in children's behavior problems. We had found intervention effects on fathers' employment. In this analysis, we also found a chain of indirect effects: in reducing personal distress and increasing co-parenting relationship quality, the intervention was also followed by increases in fathers' economic self-sufficiency.

In the primary outcome analyses presented in Section V the impact of TRUE Dads on participants was assessed for one outcome variable at a time. But family systems theory suggests that a co-parenting intervention should have reverberating effects on family functioning as a whole – on the relationships between individuals and the connections among dyads. An examination of these reverberating effects reveals some of the mechanisms by which the intervention affects the participants – how, for example, intervention-induced reductions in parents' anxiety, depression, and anger could result in reduced behavior problems in children.

We used Structural Equation Modeling to test a hypothesis based on findings from published intervention studies (Lavner, Barton, & Beach, 2020; Pruett, Cowan, Cowan, Gillette, & Pruett, 2019). We predicted that the impact of TRUE Dads on reducing parents' personal distress would set in motion a chain of effects in which couple conflict and unproductive communication would be reduced, enabling co-parents to provide more positive and less harsh parenting behavior, leading to fewer behavior problems in their children.

Structural Equation Models (SEM) are statistical techniques that enable us to trace the pathways through which interventions produce their effects. These techniques have only recently been used in family-based interventions. One advantage of these models is that instead of examining one measure at a time, they allow for the creation of multiple measures of a single construct encapsulated in what is called a latent variable. For example, instead of relying on a measure of destructive conflict as an index of couple relationship quality, we can combine destructive communication, constructive communication, constructive problem-solving style, and relationship satisfaction into a single, more comprehensive latent variable. Second, SEMs are essentially dynamic sets of multiple regressions. Simple regressions can determine whether a number of independent variables are related to a given outcome, but they cannot determine the relationships among the independent variables. SEMs have the added flexibility of being able to examine the links among all the latent variables in the equation, not simply whether ABC are related to D, but whether A is related to B, C, and D, whether B is related to C and D, and whether C is related to D. A third advantage of SEMs is that they enable us to determine whether the direct effects of an intervention (for example, on personal distress or co-parent communication) also leads to indirect effects. That is, we can determine whether there is a statistically significant pathway leading from participation in TRUE Dads through effects on different domains of the family system.

What has been relatively ignored in intervention analyses generally is that tests of direct effects do not pay attention to anything else we know about how the data may be interconnected; they simply ask: if all we know about the children of participants is whether their parents were in a program group or a control group, could we predict whether the children will have higher or lower scores on a behavior problem checklist? In this study, we know a number of additional key things about the participants that can flesh out the story. We had already shown in the primary analyses that participation in the TRUE Dads program reduced parents' symptoms of distress and destructive couple communication, although there were no *direct* effects of the intervention on parenting quality or child outcomes. Based on our previously published Supporting Father Involvement intervention studies using similar measures (Philip A. Cowan et al., 2019; Kline Pruett, Cowan, Cowan, Gillette, & Pruett, 2019) and a newly published study of other low-income families (Lavner et al., 2020), we included a latent variable representing stressful life events as a context for the stressors that affected these very low-income families.

The SEM predicted that the direct effect of the intervention on reducing parents' personal distress and increasing relationship quality would lead to less anxious or harsh parenting, with greater involvement of fathers in daily care of the child, and that, in turn, would lead to reductions in their children's behavior problems. Similarly, our model predicted that when co-

parents participated in TRUE Dads, their reduction in personal distress and increased relationship quality would facilitate fathers' progress on the road to economic self-sufficiency. We included latent variable measures of: (1) life stress events as a context factor affecting low-income families; (2) parents' personal distress (depression, anxiety, anger); (3) couple relationship quality (constructive communication, destructive communication, and satisfaction with the relationship; (4) parenting quality (parenting stress, punitive parenting, and active father engagement); (5) children's behavior problems (aggression, hyperactivity, depression, and social withdrawal); and (6) economic self-sufficiency (income, employment, confidence, and an expectation of future employment success).

We used a SmartPLS program to analyze this hypothesized SEM model, depicted in Figure 1 (Hair, Tomas, Hult, Ringle, & Sarstedt, 2016). The measures included in these constructs are described in Table IV.3. The data come from all 891 parenting teams in which at least one partner completed the one-year follow-up survey. Each latent variable includes data obtained separately from both father and co-parent, except for economic self-sufficiency, which includes data from fathers only.

We did not include every domain of our multidomain risk and protective model of change in the SEM. Although discussions of patterns in participants' family of origin were included in the curriculum, we did not expect to see intervention-related changes in this domain during the course of the study and so it is not included in the SEM. The latent variable "couple relationship quality" in the model does not include coercive violence, one of the four primary outcome variables that showed statistically significant intervention effects, because we found that the two couple relationship variables (destructive conflict and coercive violence) were highly correlated at baseline (r=.76), and their multicollinearity obscured the findings in the SEM. Including either of the two variables in the model produced the same results depicted in Figure 1.

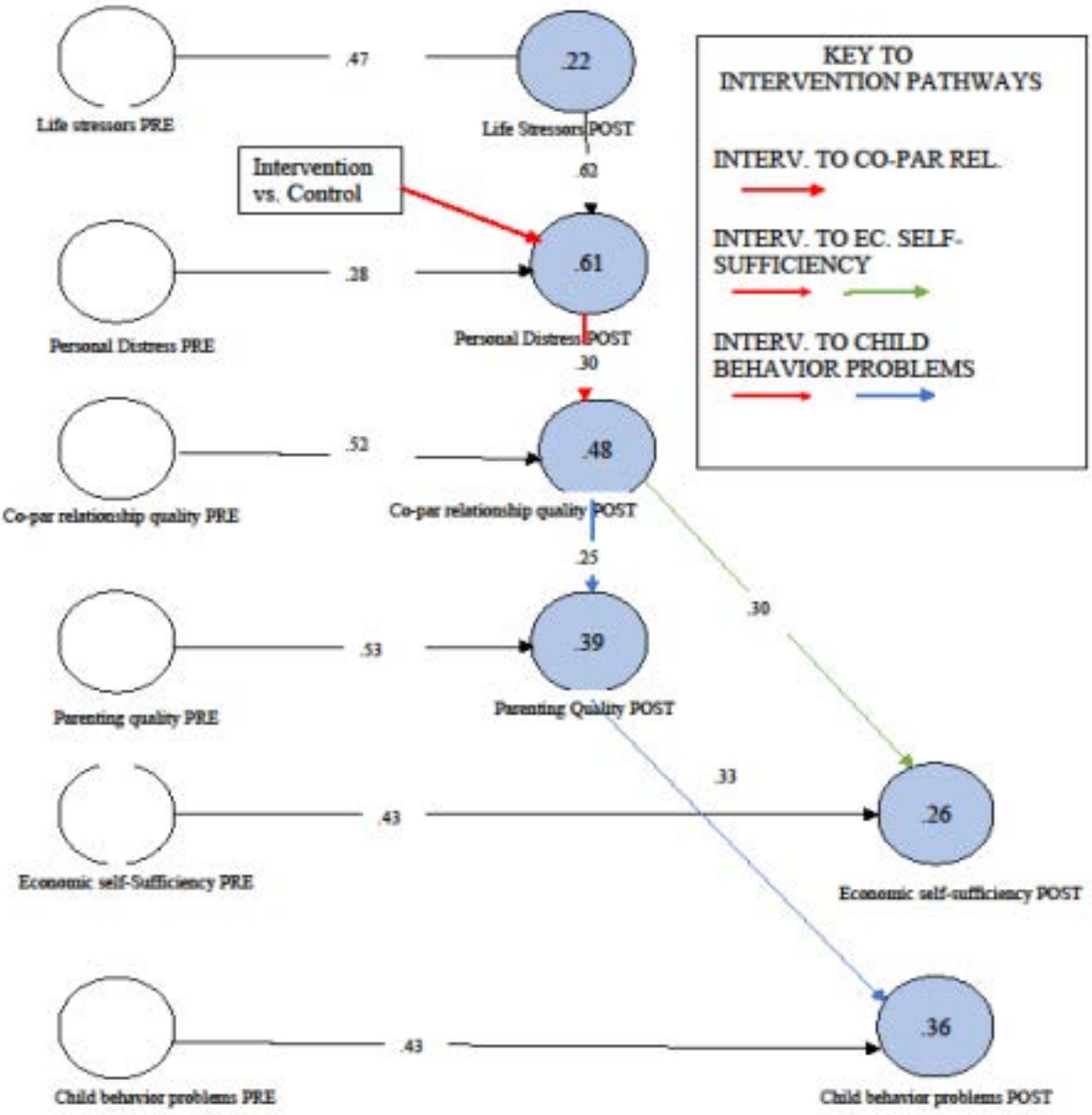
In Figure 1, the circles represent latent variables measured both at baseline (PRE) and one-year follow-up (POST), with follow-up interpreted as change from baseline. A variable representing the contrast between intervention and control participants was included in the model (the rectangular box). The numbers depicted on the lines connecting the latent variable circles are beta values describing the strength of the association. The numbers in the follow-up measures (blue circles) represent the amount of variance explained by the baseline measures plus the intervention effects.

The first key finding from the SEM model is that participation in the intervention is associated with a reduction in personal distress of both father and co-parent, and this reduction in parents' depression, anxiety, and anger leads to improvements in the quality of the co-parenting relationship (constructive and destructive communication, problem-solving style, and relationship satisfaction; t=.2.62; p=.009). A test of the reverse path, where an intervention effect on improving the quality of the co-parenting relationship might lead to a reduction in personal distress, was not statistically significant. This finding is noteworthy because couples group programs provide a curriculum heavily focused on reducing couple conflict and look for effects of the program on couple relationship quality (Knox, Cowan, Pape Cowan, & Bildner, 2011), but

the results here raise the possibility that participation in a couples group intervention has an important function of reducing the individual parent's distress, which allows them to take advantage of the communication training as a team. Because personal distress and couple relationship quality were measured at the same time, we cannot determine whether reducing personal distress is *necessary* before improvement in couple relationship quality can occur.

The second key finding is that while there are no *direct* paths linking the intervention with child behavior outcomes, there are two statistically significant *indirect* effects from parents' intervention participation to parenting quality (reductions in parenting stress and harsh parenting, increases in father involvement in the daily care of the child; t=2.19, p=.028). This improvement in parent-child relationship quality is associated with a reduction in children's behavior problems (in Figure 1, the path from red lines to blue lines; t=2.23, p=.026). In other words, the intervention does not directly affect parenting quality or children's behavior problems, but *when the parents' participation in the intervention has an impact on their personal distress*, it affects a number of other domains of the family system, as family systems theorists have long claimed.

Figure V.1. Structural Equation Model showing pathways from intervention effects to fathers' economic self-sufficiency and children's behavior problems



The third key finding is that the indirect path from intervention effects on reducing personal distress and increasing couple relationship quality leads to fathers' economic self-sufficiency (the green line: t=2.39, p=.011). Because we did find a direct effect on fathers' employment, we can be more confident in concluding that the intervention has made a difference in fathers' economic contribution to the family.

VI. Discussion

TRUE Dads began 5 years ago as a plan for an evidence-informed intervention to strengthen fathers' roles in their families. The program proposed and implemented a unique integration of fathering and co-parenting services based on the insight that the single best predictor of a father's involvement with his child is the quality of his relationship with the child's mother (Carlson et al., 2011). The present evaluation was undertaken to better understand the impact of this intervention on fathers and co-parents, their co-parent relationship and parenting behaviors, child outcomes, and fathers' economic self-sufficiency.

Our implementation analyses found that the program was offered as intended, with fidelity, although some refinements were made during the pilot phase and study. We were able to recruit 1,042 co-parenting teams composed of a father and his parenting partner. TRUE Dads was able to recruit appropriately qualified educators, case managers, and employment coaches to provide high-quality program services. Participants' attendance was close to program benchmarks for acceptable dosage. Nevertheless, encouraging attendance is one area that could be improved in future iterations of TRUE Dads. From participant surveys and focus groups, we conclude that vast majority of the participants were engaged and regarded their experience with the program positively.

The evidence presented here demonstrates that in its first systematic trial, compared with a no-treatment control group, TRUE Dads led to positive and statistically significant impacts one year after enrollment on reductions in symptoms of parents' depression, destructive communication, and control violence, and increases in fathers' employment. The effect sizes associated with TRUE Dads' impacts (.15 to .18), however, are commonly regarded as small. With regard to depressive symptoms, an effect size of .15 means that after treatment, 4.2% of the co-parenting teams randomly assigned to TD reported lower depression scores than the mean of the untreated controls. Thus, in order to have one more favorable outcome in the treatment group compared to the control group, we need to treat 25 co-parenting teams on average.

In an exploratory analysis, TRUE Dads also showed indirect effects on parenting quality and children's behavior problems using a SEM analysis. Those participants in the intervention condition who reduced their personal distress and improved their couple relationship quality became less harsh and more positively involved in parenting their children, and the children benefitted from this change. A qualification to this causal interpretation is that the measures of distress reduction and couple relationship outcomes were obtained at the same time. In order to determine whether there is actually a sequence of effects in which reduction in parents' anxiety and depression comes first, it would be necessary to have a third assessment point after the baseline and one-year follow-ups. However, the case for a causal interpretation of the findings is strengthened by the fact that the significant path linkages from intervention through personal distress, co-parenting relationship quality, parenting quality, and children's behavior problems were present in the TRUE Dads intervention sample, but not in the control group. Similar indirect effects have been found in two separate analyses of couples' group interventions (Cowan et al., 2019; Lavner et al., 2020).

The SEM yielded several important insights to supplement the impact analysis. First, while separate analyses found statistically significant impacts on depression and destructive couple communication, the SEM showed that TRUE Dads appeared to generate positive effects by providing a safe environment in which co-parents were able to reduce their depression, anxiety, and anger in ways that enabled them to adopt new and less destructive communication patterns. The findings have implications for both father involvement and healthy marriage intervention programs that use parenting teams and couples' groups as an intervention modality. Primarily, it may be important to consider anxiety-reducing strategies early in the intervention as a way of facilitating a later focus on couple or co-parenting communication.

The results showing effects on fathers' economic self-sufficiency are noteworthy for two reasons. Although the Supporting Father Involvement study, also a couples-based intervention for low-income families, found statistically significant direct impacts on family income (Cowan, Cowan, Pruett, Pruett, & Gillette, 2019), the TRUE Dads intervention is the first that we are aware of to find that participation in the intervention led to indirect effects on a latent measure of fathers' economic self-sufficiency (increases in employment, income, employment hope, and employment confidence). Second, the curricula in free-standing employment interventions are typically addressed to individuals, not couples, and focus primarily on building men's confidence and hope regarding employment, enhancing employment motivation and skills, and problem-solving to overcome employment barriers. Unfortunately, many of these programs have not proven to be successful in improving economic self-sufficiency. The SEM suggests, but does not prove, that the TRUE Dads intervention for co-parents that addressed personal, couple, parenting, and employment issues in a curriculum addressed to father and co-parent, can have a causal impact on fathers' economic contribution to the family.

These findings, then, strengthen our interpretation that intervention-induced reductions in parents' personal distress and improvements in their co-parenting relationship strategies were associated with parenting quality, the children's behavior, and the fathers' employment status in low-income families. The results suggest that TRUE Dads had the effect of helping participants understand the connections among these various aspects of family functioning. The findings also support our family systems theory of change – that intervention-induced positive changes in the parental couple relationship can have a reverberating, positive impact on other key domains of family and work life.

In the context of research as a learning process, the findings pinpoint some ways in which a revised and strengthened TRUE Dads intervention might produce stronger impacts by focusing on parents' personal distress and couple communication first, to lay fertile ground for other outcomes, so that we could ultimately also see direct effects of working with co-parents on their parenting, their children's behavior, and fathers' employment self-sufficiency. The TRUE Dads intervention demonstrates clearly that "responsible fatherhood" is a family affair in which the inclusion of co-parents in a fatherhood intervention, and improving the relationship between the partners, play a key role.

VII. References

Achatz, M., & MacAllum, C. A. (1994). *Young unwed fathers: Report from the field*. Retrieved from Philadelphia, PA:

Bollen, K.A. & Noble, M.D. (2011) Structural equation models and the quantification of behavior. *Proceedings of the National Academy of Sciences, USA. 108*, 15639-15646.

Cabrera, N. J., Volling, B. L., & Barr, R. (2018). Fathers are parents, too! Widening the lens on parenting for children's development. *Child Development Perspectives*, No Pagination Specified.

Carlson, M. J., McLanahan, S. S., & Brooks-Gunn, J. (2009). *Nonmarital fathering and the wellbeing of children.* Retrieved from Princeton, NJ:

Carlson, M. J., Pilkauskas, N. V., McLanahan, S. S., & Brooks-Gunn, J. (2011). Couples as Partners and Parents over Children's Early Years. *Journal of Marriage and the Family, 73*(2), 317-334. doi:10.1111/j.1741-3737.2010.00809.x

Clearinghouse, W. W. (2020). *What Works Clearinghouse Standards Handbook, Version 4.1. .* Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Cowan, C. P., & Cowan, P. A. (2000). *When partners become parents: the big life change for couples*. Mahwah, NJ: Lawrence Erlbaum Associates.

Cowan, C. P., & Cowan, P. A. (2018). Enhancing parenting effectiveness, fathers' involvement, couple relationship quality, and children's development: Breaking down silos in family policy making and service delivery. *Journal of Family Theory & Review, 11* 92-111. doi:DOI:10.1111/jftr.12301.

Cowan, C. P., & Cowan, P. A. (2019). Changing Families: A Preventive Intervention Perspective. *Family Relations, 68*(3), 298.

Cowan, P., Cowan, C. P., Cohen, N., Pruett, M. K., & Pruett, K. (2008). Supporting fathers engagement with their kids. In J. D. Berrick & N. Gilbert (Eds.), *Raising children: emerging needs, modern risks, and social responses* (pp. 44-81). New York: Oxford University Press.

Cowan, P. A., Cowan, C. P., & Heming, G. (2005). Five-domain models: Putting it all together. In P. A. Cowan, C. P. Cowan, J. Ablow, V. K. Johnson, & J. Measelle (Eds.), *The family context of parenting in children's adaptation to elementary school* (pp. 315-333). Mahwah, NJ: Lawrence Erlbaum Associates.

Cowan, P. A., Cowan, C. P., Pruett, M. K., & Pruett, K. (2019). Fathers' and mothers' attachment styles, couple conflict, parenting quality, and children's behavior problems: An intervention test of mediation. *Attachment & Human Development, 21*(5), 532-550.

Cowan, P. A., Cowan, C. P., Pruett, M. K., Pruett, K., & Wong, J. J. (2009). Promoting fathers' engagement with children: Preventive interventions for low-income families. *Journal of Marriage and Family, 71*(3), 663-679. doi: DOI 10.1111/j.1741-3737.2009.00625.x

Cummings, E., & Davies, P. (2010). *Marital conflict and children : an emotional security perspective.* New York: Guilford Press.

Foundation., R. W. J. (2020). County Health Rankings and Roadmap.

Gennetian, L., & Cabrera, N. J. (2018). Helping low-income fathers form healthy relationships with their children. from Institute of Family Studies https://ifstudies.org/blog/helping-low-income-fathers-form-loving-relationships-with-their-children

Hair, J. F., Jr., , Tomas, G., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Thousand Oaks, CA: Sage.

Harold, G. T., Acquah, D., Sellers, R., & Chowdry, H. E. I. F. (2016). *What Works to Enhance Inter-Parental Relationships and Improve Outcomes for Children.* Retrieved from London: http://www.eif.org.uk/publication/what-works-to-enhance-inter-parental-relationships-and-improve-outcomes-for-children-3/

Kline Pruett, M., Cowan, P. A., Cowan, C. P., Gillette, P., & Pruett, K. (2019). Supporting Father Involvement: A group intervention for low-income Community and Child Welfare referred couples. *Family Relations, 68*(1), 51-67. doi:10.1111/fare.12352

Knox, V., Cowan, P. A., Pape Cowan, C., & Bildner, E. (2011). Policies that strengthen fatherhood and family relationships: What do we know and what do we need to know? *The Annals of the American Academy of Political and Social Science, 635*(1), 216-239. doi:10.1177/0002716210394769

Lavner, J. A., Barton, A. W., & Beach, S. R. H. (2020). Direct and indirect effects of a couple-focused preventive intervention on children's outcomes: A randomized controlled trial with African American families. *Journal of Consulting and Clinical Psychology, 88*(8), 696-707.

NIMH. (2020). Men and depression. Retrieved June 29, 2020 https://www.nimh.nih.gov/health/publications/men-and-depression/index.shtml

News On 6. Oklahoma's depression rate leads the nation. May 21, 2018. Retrieved June 29, 2020 from https://www.newson6.com/story/5e35e0622f69d76f6201e4ec/oklahomas-depression-rate-leads-the-nation

Panter-Brick, C., Burgess, A., Eggerman, M., McAllister, F., Pruett, K., & Leckman, J. (2015). Practitioner Review: Engaging fathers - recommendations for a game change in parenting interventions on a systematic review of the global evidence. *Journal of Child Psychology and Psychiatry, 22*(11), 1187-1212. doi:doi:10.1111/jcpp.12280

PREP for Individuals, I. (2011). On My Shoulders: PREP's fatherhood curriculum leader manual. In. Greenwood Village, CO: PREP for Individuals.

Pruett, M. K., Cowan, P. A., Cowan, C. P., Gillette, P., & Pruett, K. D. (2019). Supporting Father Involvement: An intervention with community and Child Welfare–referred couples. *Family Relations, 68*(1), 51-67. doi:10.1111/fare.12352

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models : applications and data analysis methods* (2nd ed.). Thousand Oaks: Sage Publications.

Stanley, S. M., Markman, H. J., Jenkins, N. H., Erlacher, J., Egger, M., & Ramos, L. D. (2017). PREP 8.0 Leader Manual. In. Denver, CO: PREP Educational Products, Inc.

Statistics, B. o. L. (2018). Characteristics of minimum wage workers. https://www.bls.gov/opub/reports/minimum-wage/2018/home.htm#table10

Vieth, W. (2014). A portrait of poverty, 50 years after a declaration of 'war,'". *Oklahoma Watch*.

Whisman, M. A., & Schonbrun, Y. C. (2010). *Marital distress and relapse prevention for depression*. US: American Psychological Association.

Wood, R. G., Moore, Q., Clarkwest, A., & Killewald, A. (2014). The long-term effects of building strong families: A program for unmarried parents. *Journal of Marriage and Family, 76*(2), 446-463.

VIII. Appendices

A. Descriptive theory of change for TRUE Dads

The local evaluators of the TRUE Dads intervention have developed an empirically-based risk/protective family systems theory of change and applied it to both Couple Relationship Interventions (Cowan & Cowan, 2000; Cowan, Cowan, & Heming, 2005) and Father Involvement Interventions with low-income families (Cowan et al., 2009). This model guided the selection of topics in the TRUE Dads curriculum and the multidomain evaluation of the program. The model suggests that each of the intervention's goals (father's as well as co-parent's positive engagement with their children, positive collaboration as co-parents, and increased economic self-sufficiency) is embedded in a system of six family domains, with events in each domain potentially affecting what happens in the others (see Figure 1 in the supplementary SEM analysis for an example of data from TRUE Dads that support the theory). The six domains are:

1. History of experiences in the father's and co-parent's family of origin
2. Co-parents' physical and mental health (e.g., anxiety and depression)
3. Quality of the relationship between co-parents
4. Fathers' and mothers' ability to be effective parents
5. Attaining economic self-sufficiency
6. Managing other external stresses

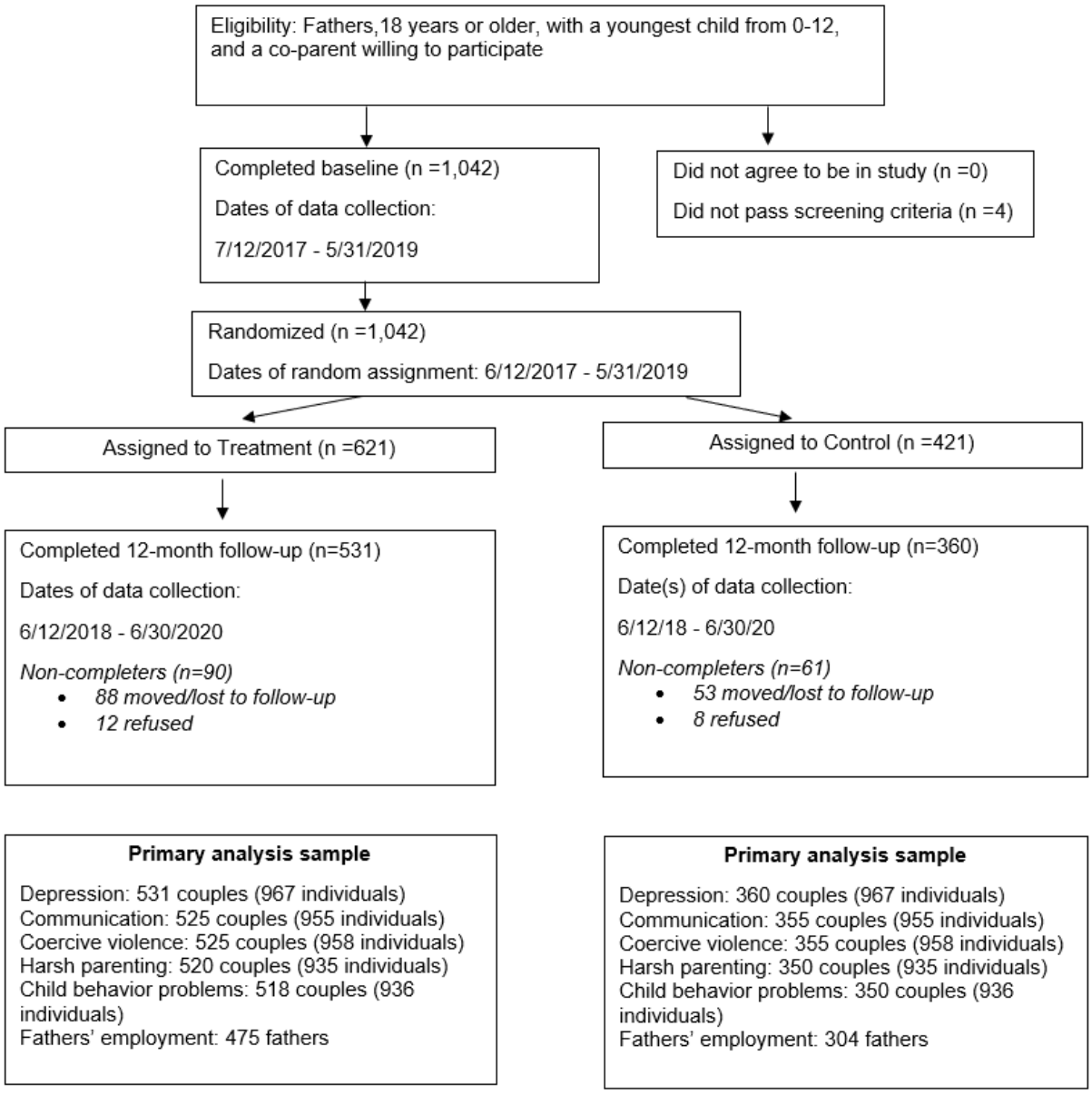
As noted in Section V, although the curriculum paid attention to family of origin issues, we chose not to include a latent variable from this domain in the SEM model because we did not expect them to change over the course of the intervention.

The TRUE Dads intervention assumes that fathers want to be involved with their children, but that services, parenting books and classes, and parenting researchers focus primarily on mothers, which sends the message that fathers are not important to their children's development. In addition, the fact that research has shown that the best predictor of fathers' involvement with their children is the quality of the father's relationship with the child's mother (Carlson et al., 2011) guided the decision to include both fathers and their co-parenting partners in the intervention.

We have cited research in the introduction to this report, along with findings from the TRUE Dads intervention, that supports the idea that intervention effects on fathers and mothers and on their relationship as co-parenting teams create pathways to improvements in parent-child relationships, children's adaptation, and fathers' economic self-sufficiency. These indirect pathways were not present in data from co-parenting teams randomly assigned to the control group.

B. CONSORT diagram for co-parenting teams

Figure A.1. CONSORT Diagram



C. Implementation research questions and analysis

We answered our research questions about program implementation using a range of data sources: quantitative data from the TRUE Dads MIS, nFORM, and surveys of program participants; reviews of personnel files; and descriptive qualitative data from case notes, staff meeting notes, and focus group interviews (See Table C.1).

- **Recruitment** questions were answered by tabulating attendance taken at each workshop session recorded in the MIS.
- **Fidelity** questions about staff education levels came from personnel files. Program refinements and modifications were captured in monthly staff meeting notes. Fidelity of program delivery was evaluated by periodic observation and feedback by staff supervisors. Because educators spent time working with material elicited from the participants in response to the curriculum, we did not feel that simple checklists of material covered were adequate to establish fidelity to TRUE Dad's overall conception.
- Participant **dosage** was established based on workshop attendance recorded in the MIS.
- Answers to questions about program **quality** relevant to staff training and supervision included data from training schedules, supervisor notes, program manager notes and emails, and minutes of regularly scheduled conference calls between Oklahoma staff and program evaluators.
- Participant ratings of program **quality** and **engagement** were answered with both quantitative and qualitative data. The quantitative data came from averaged responses to a number of questions asked at the one-year follow-up. Qualitative data were obtained in focus group interviews by Prof. Sarah Halpern-Meekin. The quantitative ratings of participants' experience included items that reflected their engagement with the group leaders and other group members and their ratings of how much they learned. Those participants who attended the focus groups were extremely enthusiastic about their experience. Of course, because of its construction, the focus group responses were limited to those who completed the intervention and were willing to attend a focus group.

Table C.1. Source of data for analysis of implementation questions

Implementation element	Research question	Data source	Timing/frequency of data collection	Party responsible for data collection
Recruitment	• Can we recruit co-parenting teams to TRUE Dads?	IMCI MIS, nFORM, Qualtrics	Intake data	Intake staff
	• Will participants accept random assignment?	IMCI MIS	Intake data	Intake staff
Fidelity	• Did the Educators have adequate educational background to deliver this program?	Personnel files	Applications for facilitator position	Intervention staff
	• Steps to modify/improve TRUE Dads during the pilot phase?	Program Manager's notation of regular staff meetings to review how the program was going	Approximately once per month	Program Manager
	• Did the Educators deliver the intended curriculum content to the participants?	Educator checklist after each session Occasional observations	Every group session Approximately once per month	Educators Supervisor

Implementation element	Research question	Data source	Timing/frequency of data collection	Party responsible for data collection
Dosage	• How often did the intervention group participants attend the workshops on average? • Did attendance reach benchmark goals?	Workshop sessions and individual service contacts in MIS; attendance logs	All sessions took attendance	Intervention staff
Quality	• What mandatory training did the Educators have?	Program Manager notes	One-time collection of program manager notes	Program manager
	• What steps were taken (for example, supervision or additional training) to improve the quality of service delivery?	Staff meetings Further training Individual supervision	Bi-monthly Every 6 months Monthly	These were scheduled events without data collection
	• Was the program perceived positively by the participants?	Qualitative focus groups Qualtrics follow-up items	End of years 2 and 4 1 year after participants entered the study	Sarah Halpern-Meekin, Ph.D. Program staff
Engagement	• Were participants actively engaged in the program?	Qualtrics follow-up items Written comments at the end of Core and Intensive workshops	1 year after participants entered the study End of 6th workshop and end of 12th workshop	Program staff
Context	• What other similar HM&RF programs were available to control group members?	Survey of local and State family service resources	Ongoing	Program staff

IMCI = It's My Community Initiative; MIS = Management information system; nFORM = Information, Family Outcomes, Reporting and Management

D. Data preparation

The program staff sent study data to the evaluators at the University of California, Berkeley for analysis in digital form, with identifying data removed and replaced with participant code numbers. Study participants or staff entered some basic personal and demographic data on iPads at intake, and these data were stored on the IMCI MIS. Participants entered nFORM Applicant Characteristics and Entrance and Exit surveys on iPads. Participants entered Qualtrics survey data on iPads at Intake interviews and again at follow-up one year later. The Qualtrics data included only code numbers that allowed us to match these files with MIS file and nFORM data.

The Qualtrics data were cleaned by identifying participant data entry errors (e.g., numbers outside the range of the item, entering age of father in the child age item). Errors were checked against data from the MIS and resolved. Where no replacement information was available, the data were recorded as missing.

We identified a number of items across respondents that allowed us to check whether both partners in the co-parenting team were giving the same responses to items where we would

expect to see agreement (e.g., age of child, marital status). For items such as marital status, we select the less publicly committed response as close to the correct answer (e.g., cohabiting, if one parent checked "married" and the other checked "living together"). For child's age, we selected the response of the co-parent if an intimate partner. For items that reflected each partner's perception or opinion (e.g., couple relationship satisfaction, child's level of aggressive behavior), we included both partners' responses regardless of their difference.

We checked for missing items within scales and found no scale with more than 5% of the items missing in the overall sample. We created scales in two ways: one with the sum of the items, which then included missing data, and one with the mean of the items present, which in essence, imputed the missing value as the mean of the remaining items. This was allowed if 20% or fewer of a scale's items were missing for a participant's response. The two scores were so highly correlated (above .95 in all cases) that we selected the mean score for the analytic data set.

We determined whether our measures for the primary impact analysis had acceptable reliability (Cronbach's alpha > .70) using item reliability statistics produced in SPSS. All the measures were reliable. We determined whether our latent variable measures had items significantly correlated with the latent construct using the SmartPls program and found that all the items were significantly correlated with their intended construct.

E. Impact estimation

TRUE Dads is a couples-based intervention. Although it was possible to analyze the data from fathers and their co-parenting partners separately, both theory and the empirical findings of this study suggest that the data from the partners are interdependent. Table E.1 shows the low to moderate correlations between partners for all 5 outcome measures.

Table E.1. Correlations between father and co-parent on primary outcome measures at baseline

Measure	r
Depression	.165
Destructive communication	.482
Coercive violence	.447
Harsh parenting	.350
Child behavior problems	.278

The question, then, was how to construct a data set that would allow us to perform couple-level analyses, especially when only one of the partners completed the 1-year follow-up survey.

We considered four alternatives to handling missing data and creating a data set for analysis:

1. Limiting the data set to co-parenting teams in which both partners complete the 1-year follow-up survey: In our view this approach introduces serious bias in the comparison between TRUE Dads and control participants. On the average (See Table V.3) there were 693 teams with *both partners* completing the follow-up and 891 teams with *either partner* completing. To include only co-parenting teams in which both completed the follow-up would mean that we would be reducing statistical power and "throwing out" data from 184

teams – that is, from the couple perspective, 21% of the co-parenting teams. We believe that this approach is unacceptable.

2. Using traditional Analysis of Covariance (ANCOVA) methods: In the present study a traditional ANCOVA analysis would treat data from fathers and co-parents as related measures and enter fathers' and co-parents' data for each measure at follow-up, with fathers' and co-parents' data at baseline as covariates. The analysis would yield an intervention effect and a father/co-parent intervention interaction effect. The problem with this approach is that SPSS and other data analysis programs applying ANCOVA analyses discard any cases with missing data, so that, in effect, the data set would be limited to cases in which both partners complete the 1-year follow-up. This approach results in the same problem as alternative 1.
3. Creating couple-level variables by averaging scale scores from both partners on each measure: If one partner's follow-up data are missing, then the couple score is represented by the other partner. With this approach, it is possible to perform ANCOVA analyses to determine if the intervention affects the couple outcome. This solution has been adopted by previous ACF-funded large-scale Healthy Marriage projects – Building Strong Families (Wood, Moore, Clarkwest, & Killewald, 2014) and Strengthening Healthy Marriage (Hsueh et al., 2012).
4. Conducting Hierarchical Linear Modeling (HLM) analysis, which uses all data and examines both individual and couple trajectories over time (Raudenbush & Bryk, 2002): At first glance this approach seems to solve the missing data problems outlined above and avoids averaging co-parents' responses. An exploratory re-analysis of the data from the TRUE Dads project showed that the correlation between effect size estimates obtained from the ANCOVA analyses in alternative 3 and those obtained with HLM analyses was r=.85 (p<.01).

Approaches 3 and 4 arrive at similar conclusions, but the ANCOVA in approach 3 produces corrected means and is easier to interpret. Given the simplicity of ANCOVAs in reporting the means and standard deviations for intervention and control samples, we performed the analyses of intervention impact using ANCOVAs based on mean co-parenting team scores. For the primary impact analysis, the impact of participating in the intervention was determined by comparing the difference between intervention and control means at the one-year follow-up, with scores at baseline for the outcome measure entered as covariates.

The statistical model for Analysis of Covariance is:

$$Y = B_0 + B_1 X_1 + B_2 X_2 + e$$

Where:

Y = the outcome of interest at follow-up
B_0 = the Y intercept
B_1 = the difference between the intervention and control groups (the effect of the intervention on the outcome of interest)
B_2 = the baseline measure of the outcome variable
e = error term

Baseline equivalence. Baseline differences and *p*-values for dichotomous data were calculated with a Chi-square test, and effect size estimates were calculated using the Cox index. Baseline differences and *p*-values for continuous data were calculated with *t*-tests, and effect sizes estimated using Hedges-*g*. Details of the rationale for our baseline equivalence analysis can be found in Appendix F.

F. Baseline equivalence of analytic samples

Using the analytic data set relevant to destructive communication and coercive violence (n=880; See Table F.1), we found only one statistically significant baseline difference: the intervention group had a higher proportion of parenting-teams who were not intimate partners. The effect size (Cox *d)* was .20. In addition, the slightly lower proportion of Hispanic participants in the intervention condition at baseline compared with the controls was not statistically significant but had an effect size of *d* =.15, and the non-statistically-significant baseline differences in the co-parent having a high school diploma had an effect size of *d*=.15.

Table F.1. Key baseline measures and baseline equivalence across study groups, for 880 co-parenting teams who answered questions about destructive communication and control violence at one-year

Baseline measure	Intervention mean (standard deviation)	Comparison mean (standard deviation)	Intervention versus comparison mean difference (p-value of difference)	Effect size
Fathers (%)	50	50	0	0
Fathers' Race/ethnicity (%)				
Hispanic	9.0	11.2	-2.2 (*p*=.175)	.15
Native American	13.7	13.8	-.1 (*p*=.956)	.01
Non-Hispanic White	28.0	26.6	1.4 (*p*=.585)	.04
Non-Hispanic Black	43.2	41.7	1.5 (*p*=.578)	.04
Married (%)	34.4	31.2	3.2 (*p*=.317)	.09
Living together (%)	84.7	83.9	.8 (*p*=.767)	.04
Non-intimate partner (%)	17.4	13.2	4.2 (*p*=.022)*	.20
Father is unemployed (%)	38.9	37.0	1.9 (*p*=.567)	.05
Father has high school diploma or above (%)	63.8	61.2	.6 (*p*=.457)	.07
Co-parent has high school diploma or above (%)	70.2	75.4	-5.2 (*p*=.112)	.15
Father's annual income ($)	11,500	10,250	1,250 (*p*=.476)	.05
Co-parent's annual income ($)	8,600	7,750	850 (*p*=.386)	.06
Father's age	31.7 (7.61)	31.8 (7.58)	-.1 (*p*=.852)	.01
Co-parent's age	31.0 (8.19)	30.9 (8.59)	.1 (*p*=.837)	.01
Youngest Child's age	3.0 (3.14)	2.8 (3.04)	.2 (*p*=.292	.10

Baseline measure	Intervention mean (standard deviation)	Comparison mean (standard deviation)	Intervention versus comparison mean difference (p-value of difference)	Effect size
Baseline scores for the primary outcome variables for co-parenting teams in which either partner completed 1-year follow-up				
Destructive communication[1]	1.35 (.79)	1.37 (.77)	-.02 (*p*=.707)	.03
Coercive violence[1]	.26 (.23)	.26 (.23)	0.0	.00

[1]Intervention n=525; control n=355

*p<.05

Finally, using the analytic sample relevant to harsh parenting and child behavior problems (n-808; See Table F.2) we found one statistically significant baseline difference between intervention and controls. In this analysis, the difference in proportion of Hispanic participants was statistically significant with an effect size of .25, with only one other effect size (for slightly more control co-parents having a high school diploma or above), but the difference was not statistically significant.

Table F.2. Key baseline measures and baseline equivalence across study groups, for 808 co-parenting teams who answered questions about harsh parenting and child behavior problems at one-year

Baseline measure	Intervention mean (standard deviation)	Comparison mean (standard deviation)	Intervention versus comparison mean difference (p-value of difference)	Effect size
Fathers (%)	50	50	0	0
Father' Race/Ethnicity (%)				
Hispanic	8.4	11.2	-2.8 (*p*=.040)*	*.25*
Native American	13.8	13.5	0.3 (*p*=.807)	.04
Non-Hispanic White	27.8	25.7	2.1 (*p*=.743)	.03
Non-Hispanic Black	43.7	43.1	0.6 (*p*=.409)	.08
Married (%)	34.6	31	3.6 (p=.289)	.10
Living together (%)	84.7	83.3	1.4 (*p*=.637)	.06
Non-intimate partner (%)	17.0	12.9	4.1 (*p*=.580)	.08
Father is unemployed (%)	39.3	36.4	2.9 (*p*=.394)	.08
Father has high school diploma or above (%)	63.1	61.6	1.5 (*p*=.602)	.06
Co-parent has high school diploma or above (%)	71.0	74.8	-3.8 (*p*=.260)	.12
Father's annual income ($)	11,350	9,750	1,600 (*p*=.167)	.10
Co-parent's annual income ($)	8,650	7,900	700 (*p*=.445)	.05
Father's age	31.94 (7.52)	31.97 (7.79)	-.03 (*p*=.546)	.00
Co-parent's age	31.05 (8.83)	31.18 (8.20)	-.13 (*p*=.950)	.02

Baseline measure	Intervention mean (standard deviation)	Comparison mean (standard deviation)	Intervention versus comparison mean difference (p-value of difference)	Effect size
Youngest Child's age	3.17 (3.15)	3.04 (3.06)	.13 (*p*=.546)	.05
Baseline scores for the primary outcome variables for co-parenting teams in which either partner completed 1-year follow-up				
Harsh parenting[1]	.74 (.71)	.68 (.69)	.06 (*p*=.248)	.09
Child behavior problems[1]	1.75 (.39)	1.73 (.38)	.01 (*p*=.303)	.05

[1]Intervention n=490; Control n=318
*$p<.05$

These baseline equivalent analyses suggested that the randomization process worked well, and that the Benchmark sample and these two smaller samples resulted in all outcome measures and almost all demographic measures being equivalent. Nevertheless, these analyses suggested that there were a few demographic measures that should be included in sensitivity analyses (See Tables V.3 and V.4).

G. Additional analyses

Path models are made up of the **structural model**, describing the relationships among the latent variables, and the **measurement model**, describing the correlation of each of the measures of construct with the latent variable (Hair et al., 2017).

The following is adapted from Bollen & Noble (2011):

The *structural model* is (Eq. **1**)

$\eta_i = \alpha_\eta + \mathbf{\Gamma}\ \xi_i + e$ where η_i is a latent dependent variable i (e.g., child behavior problems), α_η is the intercept of dependent variable i, ξ_i is the vector of latent independent variables (e.g., intervention, personal distress, etc.), $\mathbf{\Gamma}$ is the coefficient matrix giving the effects of the latent independent variables on the latent dependent variables (e.g., on child behavior problems or economic self-sufficiency) and e is the error of measurement.

The latent variable model reflects the hypotheses that the different constructs are connected to each other (for example, the relation between parents' personal distress and the quality of the co-parenting relationship).

The *measurement mod*el links the latent variable (the statistically determined index representing the construct) to the various measures of the construct (the measures included "inside" each of the circles). It has two equations (Eqs. **2** and **3**),

$$\mathbf{y}_i = \alpha_y + \Lambda_y \eta_i + \varepsilon_i \text{ and } \quad [2] \quad \mathbf{x}_i = \alpha_x + \Lambda_x \xi_i + \delta_i, \quad [3]$$

where $\mathbf{y}_i$ and $\mathbf{x}_i$ are vectors of the observed indicators of η_i and ξ_i, respectively, α_y and α_x are intercepts, $\mathbf{\Lambda}_y$ and $\mathbf{\Lambda}_x$ are factor loadings or regression coefficients that give the relative weight assigned to each measure in forming the construct.