ROBERT G. WOOD AND QUINN MOORE   *Mathematica Policy Research*

ANDREW CLARKWEST   *Mathematica Policy Research**

ALEXANDRA KILLEWALD   *Harvard University***

The Long-Term Effects of Building Strong Families: A Program for Unmarried Parents

The authors present findings from a large-scale, random-assignment evaluation of Building Strong Families (BSF), a program offering group sessions on relationship skills education to low-income, unmarried parents who were expecting or had recently had a baby. Findings based on a 3-year follow-up survey of over 4,000 couples indicate that BSF did not succeed in its central objectives of improving the couple relationship, increasing the quality of coparenting, or enhancing father involvement. In fact, the program had modest negative effects on some of these outcomes. BSF also had little impact on child well-being, with no effect on children's family stability or economic well-being and only a modest positive effect on children's socioemotional development. Impacts varied across the 8 study sites. Although attendance at group sessions was relatively low, there is little evidence of program effects even among couples who attended sessions regularly.

Mathematica Policy Research, P.O. Box 2393, Princeton, NJ 08543-2393 (rwood@mathematica-mpr.com).

*Mathematica Policy Research, 1100 1st Street, NE, 12th Floor, Washington, DC 20002–4221.

**Department of Sociology, Harvard University, 648 William James Hall, 33 Kirkland St., Cambridge, MA 02138.

In 2010, over 40% of U.S. births were to unmarried parents (Martin et al., 2012). Most of the parents were in romantic relationships when their children were born; about half were living together (Carlson, McLanahan, England, & Devaney, 2005; Sigle-Rushton & McLanahan, 2002). Unmarried parents are typically optimistic about their future together, including the likelihood that they will eventually marry (Carlson et al., 2005). But these hopes are often unrealized. Most are no longer in a romantic relationship 5 years after their child's birth (Center for Research on Child Wellbeing, 2007).

The disruption of unmarried parents' relationships is significant because children growing up in households that do not include both their biological parents are at greater risk of poor outcomes (Amato, 2005; Carlson & Corcoran, 2001; Kim, 2011; McLanahan & Sandefur, 1994). For this reason, there is considerable interest among policymakers in supporting unmarried parents' efforts to maintain their relationships. The 1996 Personal Responsibility and Work Opportunity Reconciliation Act made increasing the number of children raised in two-parent families an explicit policy objective. In 2001, the Administration for Children and Families (ACF), U.S. Department of Health and Human Services, launched the Healthy Marriage Initiative. In 2005, Congress approved $100 million in annual funding for programs designed to encourage and strengthen marriage. This funding was continued in 2010

when Congress voted to provide $75 million annually for marriage education.

Building Strong Families (BSF) was one of the first major projects under ACF's Healthy Marriage Initiative. The project developed, implemented, and tested voluntary programs designed to help unmarried, economically disadvantaged parents who were expecting or had just had a baby strengthen their couple relationship, with the ultimate goal of helping them create a stable and healthy home environment for their children. The main component of the BSF model was relationship skills education offered to couples in group sessions. The model was implemented in eight sites and tested using a random-assignment research design involving over 5,000 couples. To track couples' outcomes, telephone surveys were conducted with both mothers and fathers 15 and 36 months after they applied for BSF.

Results based on the 15-month follow-up were released in 2010 (Wood, McConnell, Moore, Clarkwest, & Hsueh, 2010). That analysis, which focused on interim outcomes— primarily the couples' relationship status and quality—found that BSF had no effect on these relationships when results from the eight evaluation sites were combined. However, the results varied across the eight sites, with one program having a consistent pattern of positive effects and another having a number of negative effects. The other six programs generally had little or no effect on relationships.

In this article, we present final BSF impact results based on the 3-year follow-up. The analysis examines how the somewhat mixed picture that emerged at 15 months has evolved over the longer term. It is important to examine long-term impacts because changes in relationship status can take time to unfold. Moreover, in this article we extend the analysis to additional domains associated with child well-being that were not examined in the 15-month analysis. The ultimate aim of BSF was to improve child well-being by improving parents' relationship stability and quality. To examine BSF's success in this area, we estimated impacts on the well-being of the children who made these couples eligible for BSF services: those born around the time couples applied for the program and who were about 3 years old at the final follow-up.

In addition, we examined BSF's effects on the couples who actually received the program's core service of relationship skills education. Getting couples to attend sessions proved challenging, and 45% of couples who were offered program services never attended these sessions. Because of this large proportion of nonattending couples, we used quasi-experimental techniques to examine BSF's effects on two groups: (a) couples who attended any group sessions and (b) couples who attended at least half of these sessions.

BACKGROUND

Related Literature

Relationship skills education programs have existed for several decades. Evidence has suggested that these programs can be effective at improving the relationship quality of middle-class, married couples (Dion, 2005; Hawkins, Blanchard, Baldwin, & Fawcett, 2008; Markman, Renick, Floyd, Stanley, & Clements, 1993; Reardon-Anderson, Stagner, Macomber, & Murray, 2005). Until recently, little rigorous research existed on the effectiveness of similar programs for low-income couples or unmarried parents, and few studies have examined long-term impacts of these programs. A 2010 meta-analysis of research on relationship skills programs serving low-income couples, both married and unmarried, found generally positive impacts. The authors noted that the research summarized was typically not rigorous. Most studies used pre–post designs with no comparison group, sample sizes were small, and outcomes were typically measured shortly after the end of the program (Hawkins & Fackrell, 2010).

Several recent studies of relationship skills programs serving low- and moderate-income couples have used experimental research designs. A study of the Prevention and Relationship Enhancement Program (PREP) for Strong Bonds, a relationship skills education program serving married military couples, found that the program lowered the risk of divorce during the 2 years after the end of the program; no impacts were found on self-reported relationship quality (Stanley et al., 2013). The Fatherhood, Relationship, and Marriage Education project tested a version of the PREP for Strong Bonds curriculum adapted for low-income parents and found positive effects on conflict management measured 2 weeks after the program ended

(Wadsworth et al., 2011). The Supporting Father Involvement study examined the effectiveness of a curriculum delivered to low-income couples in group sessions. The Supporting Father Involvement program focused primarily on increasing fathers' involvement with their children but also aimed to improve the couple relationship. The authors reported positive impacts on relationship satisfaction at the 18-month follow-up (Cowan, Cowan, Pruett, Pruett, & Wong, 2009). The ACF-sponsored Supporting Healthy Marriage (SHM) evaluation tested a program similar to BSF but serving low-income married couples. The study found relatively small positive effects on couple's relationship quality at the 12-month follow-up, but no effect on relationship stability (Hsueh et al., 2012). These studies included primarily—and, in the case of both SHM and PREP for Strong Bonds, exclusively—married couples.

Unmarried parents face distinct relationship challenges that could influence the effectiveness of relationship skills education programs. These challenges include lower levels of commitment, higher rates of infidelity, more frequent multiple-partner fertility, and greater economic vulnerability (Carlson & Furstenberg, 2006; Edin, England, & Linnenberg, 2003). Very little rigorous research has examined the effectiveness of relationship skills education programs serving primarily unmarried parents. The Young Parenthood Program, an 8- to 12-week intervention designed to improve the relationship skills of coparenting adolescents, was evaluated with a random assignment research design. The study found short-term, marginally statistically significant reductions in intimate partner violence that did not persist at the final follow-up 18 months after the child's birth (Florsheim, McArthur, Hudak, Heavin, & Burrow-Sanchez, 2011). To date, BSF is the only random-assignment evaluation of a relationship skills education program serving primarily unmarried adult parents. In addition, it is one of only a handful of studies of relationship skills programs that used a large research sample and among the few that has a follow-up period extending 3 years after study enrollment. Thus, the analysis of the 3-year BSF follow-up data presented here provides important new information on the long-term effects of offering relationship skills education to low-income, unmarried parents.

The BSF Program

BSF was a voluntary program for new unmarried parents. Couples were eligible to enroll if they had had a baby in the last 3 months or were expecting one, were both at least 18 years old, and were romantically involved and unmarried at the time their baby was conceived. Eight organizations were selected to implement the BSF program model, which laid out a set of research-based program guidelines (Hershey, Devaney, Dion, & McConnell, 2004). Four sites (those in Florida, Indiana, and Houston and San Angelo, Texas) added BSF services to their Healthy Families programs, which aim to promote positive parenting via staff visiting and educating new and expectant parents in their homes. Other programs (those in Baltimore, Maryland, and Baton Rouge, Louisiana) were offered by agencies that provided a mix of services to low-income families. The infrastructure for programs in Atlanta, Georgia, and Oklahoma City, Oklahoma were built from the ground up specifically for BSF. Across the eight programs, key recruitment sources included hospital maternity wards, prenatal clinics, health clinics, and clinics for the Special Nutritional Program for Women, Infants, and Children.

The core component of BSF was curriculum-based group education on relationship skills. Programs chose one of three curricula that had been adapted by the developers to the needs of unmarried parents. These adaptations were aimed to better address issues specific to low-income unmarried couples and enhance the cultural sensitivity of the intervention. For example, early focus groups revealed that many potential participants had negative interactions with educational institutions and wished to avoid being lectured. In response, curricula were modified to use a less didactic approach and to allow couples to share their own experiences and learn from each other. In addition, topics were added to the curricula that research suggests are particularly important for low-income, unmarried couples, including skills for building trust and commitment and managing relationship issues related to multiple-partner fertility (McConnell et al., 2006).

The program in Oklahoma City chose the "Becoming Parents for Low-Income and Low-Literacy Couples" curriculum, developed by Pamela Jordan. The San Angelo and Houston programs chose "Love's Cradle," developed

by Mary Ortwein and Bernard Guerney. The other five programs chose "Loving Couples, Loving Children," developed by John and Julie Gottman. These curricula covered a standard set of topics. However, the emphasis varied somewhat, with "Becoming Parents" emphasizing the transition to parenthood. In addition, "Becoming Parents" included only 30 hours of material, compared with 42 hours for the other two curricula, and could thus be delivered in fewer sessions. Under all three curricula, group sessions usually met weekly and ranged in length from 2 to 5 hours, with shorter sessions typically held on weeknights and longer sessions held on weekends. Depending on the format and the number of hours of instruction offered, the curriculum could take as little as 6 weeks or as much as 5 months to complete. The sessions were led by facilitators who were trained directly by the curriculum developers. The developers also provided ongoing technical assistance in an effort to promote fidelity of program implementation (Dion et al., 2008).

BSF complemented the group relationship skills education with other supports, including a family coordinator who reinforced relationship skills, provided emotional support, and encouraged participation in the group sessions. The family coordinator also assessed family members' needs and referred them for appropriate support services, such as education, employment, and mental health services. In the four Healthy Families programs that adopted BSF programs, home visitors were assigned to fill the BSF family coordinator role and continued providing Healthy Families services during home visits.

The BSF evaluation included an extensive implementation study of the delivery of program services in the eight BSF evaluation sites (Dion et al., 2008; Dion, Avellar, & Clary, 2010). The study found that the eight sites successfully implemented BSF's core components as specified by the program model. In addition, program participants in focus groups and in-depth interviews reported benefiting from the communication and conflict management techniques taught in group sessions as well as the social network offered by the groups (Dion et al., 2010). Despite these positive views of program services, the implementation study also revealed that sites had difficulty getting couples to attend group sessions regularly.

Programs devoted substantial effort to promoting attendance at group sessions, including free meals, on-site child care, and transportation assistance. Some programs also offered cash incentives, gift cards, or baby products to promote attendance. Despite these efforts, across the eight programs, only 55% of the couples assigned to the program group attended a group session. Couples who attended at least one session typically attended multiple sessions and received a substantial dose of this program component, averaging 21 hours of attendance. In a later section of this article we examine impacts for this group of couples as well as those who attended at least half of the sessions. Most couples who did not attend group sessions received other services from the program, such as help from a family coordinator or referrals to support services. Overall, 90% of couples who enrolled in BSF received some service from the program.

Analytic Approach

The main pathway by which BSF was expected to improve the lives of participating families was by improving the quality of the couple's romantic relationship. The hope was that improved relationship quality would increase the likelihood that couples remained together in a healthy relationship, which would ultimately improve child well-being. It was also thought that the program could improve coparenting and father involvement either directly or indirectly through its effects on the couple relationship. BSF's potential effects on coparenting and father involvement could in turn contribute to the program's intended influence on child well-being.

The BSF intervention has the potential to affect multiple aspects of couples' lives. Examining a large number of outcomes in an impact analysis increases the risk of finding a statistically significant result by chance (Schochet, 2009). To address this multiple-comparison concern, our analysis focused only on the outcome domains BSF was intended to influence most directly and made use of a relatively small set of outcomes within each outcome domain.

We organized our analysis around the three broad areas that BSF hoped to affect: (a) the couple relationship, measured by relationship quality and relationship status; (b) coparenting

and father involvement; and (c) child well-being, measured by family stability, economic well-being, and socioemotional development. An examination of impacts on these outcomes served as a test of whether the program succeeded in its primary objectives of improving couples' relationships, their parenting, and their children's well-being.

METHOD

Study Design and Research Sample

Couples who applied to the BSF program were assigned randomly to either the BSF group that was offered admission to the program or to a control group that was not. Across the eight programs, 5,102 couples were randomly assigned from July 2005 to March 2008, with 2,553 assigned to the BSF group and 2,549 assigned to the control group. As illustrated in Table 1, random assignment created two research groups (BSF couples and control group couples) with very similar characteristics at baseline.

Most couples in the research sample were in stable relationships and aspired to marriage when they applied for the program. Across all programs, 7% of couples were married at program application, having wed after their baby was conceived but before applying for BSF (see Table 1). About 6 in 10 were unmarried and reported living together all of the time. Among couples unmarried at application, almost 60% reported that they both thought there was either "a pretty good" or "an almost certain" chance that they would marry each other.

The couples who applied for BSF faced many potential stresses to their relationships. Their earnings were generally low, with the combined earnings of the mother and father averaging about $20,000 in the year prior to BSF application. Only 37% of couples included two members with high school diplomas. In nearly half of all couples applying for BSF, at least one of the parents had a child from a prior relationship.

BSF served a racially and ethnically diverse population. About half the couples were African American, about one quarter were Hispanic, and 12% were White (see Table 1). The rest were interracial couples or couples in which both parents considered themselves neither White, nor African American, nor Hispanic.

Data Sources

We used data from two sources: (a) a baseline form completed by all parents when they applied to BSF and (b) telephone surveys conducted about 3 years after program application. In both research groups, 80% of mothers and 69% of fathers responded to the 36-month follow-up survey. At least one parent responded in 4,247 couples (85% of all couples). This response rate was the same for the two research groups. We used the sample of couples for which at least one member responded to estimate most program impacts. If only one member responded, we used a multiple-imputation strategy to impute the response of the other partner.

Measures Related to the Couple's Relationship

Relationship status. We examined impacts on three measures of the couple's relationship status at follow-up: (a) whether both members of the couple reported being romantically involved, (b) whether both reported living together (either married or unmarried) all or most of the time, and (c) whether both reported being married to each other. When couple members disagreed about their relationship status, we used a simple rule: A couple was categorized as having a particular status only if both members of the couple reported that status. When there was a discrepancy between the two responses, the couple was assigned to the "no" category for that particular outcome.

In the large majority of couples, the partners agreed on their relationship status. The rate of disagreement between partners was 9% for romantic involvement, 9% for coresidence, and 2% for marital status. Similarly high rates of partner disagreement on relationship status among unmarried parents were found in the nationally representative data from the Fragile Families and Child Well-Being Study (Knab & McLanahan, 2007). For all relationship status measures, rates of mother–father disagreement were similar in the two research groups.

Relationship quality. Using the 37 relationship quality questions from the follow-up survey, we conducted a factor analysis to develop a small set of key relationship quality measures for the impact analysis. Five distinct relationship quality measures emerged from this analysis:

Table 1. *Baseline Characteristics of the Building Strong Families (BSF) Program and Control Group Couples*

Characteristic	BSF couples	Control group couples
Socioeconomic and demographic characteristics		
Race/ethnicity (%)		
Both partners are Hispanic	25.7	25.4
Both partners are Black, non-Hispanic	47.5	46.8
Both partners are White, non-Hispanic	11.5	11.5
All other couples	15.4	16.2
High school diploma receipt (excluding GEDs; %)		
Both partners have diploma	36.4	37.8
One partner has diploma	36.9	36.3
Neither partner has diploma	26.8	25.9
Average age (years)		
Mother's age	23.6	23.5
Father's age	26.0	25.8
Couples' total earnings in past year	$20,651	$19,866*
Either partner received TANF or SNAP in past year (%)	46.0	45.2
Relationship characteristics		
Couple's relationship status (%)		
Married to each other	6.6	7.0
Unmarried, cohabiting full time	59.9	57.3*
Unmarried, not cohabiting full time	33.5	35.7
Relationship quality[a]		
Highest tercile (%)	31.6	32.8
Middle tercile (%)	35.6	33.4
Lowest tercile (%)	32.9	33.9
Average scale value (range: 1–4)	3.26	3.25
Both partners expect to marry (%)	59.5	57.8
Baby born prior to BSF application (%)	43.9	43.8
Either partner has a child from a prior relationship (%)	47.6	46.7
Pregnancy intendedness (%)		
Intended by both partners	25.0	24.2
Wanted by both partners, but considered mistimed	52.8	53.1
Unwanted by at least one partner	22.2	22.7
Mental health, attitudes, and religiosity		
Either partner has psychological distress[b] (%)	38.0	38.3
Both partners agree with the statement "It is better for children if parents are married" (%)	61.2	59.9
Attendance at religious services (%)		
Both attend more than monthly	24.5	23.6
One attends more than monthly	28.5	28.9
Neither attends more than monthly	47.0	47.5
Sample size	2,553	2,549

Note: Data are from BSF baseline information forms and BSF eligibility forms. The eight BSF programs are weighted equally for these calculations. TANF = Temporary Assistance for Needy Families; SNAP = Supplemental Nutrition Assistance Program.

[a]This scale is based on nine items asking respondents to rate on a 4-point scale, which ranged from *strongly disagree* to *strongly agree*, the extent to which their partner (1) shows love and affection, (2) gives encouragement, and (3) listens; (4) respondents' satisfaction with how the couple resolves conflict; (5) whether the couple enjoys doing things together; respondents' (6) marriage expectations, (7) confidence in partner's fidelity, (8) confidence in wanting to be with partner in the future, and (9) feeling that the relationship with their partner is the most important thing to them. [b]Psychological distress was assessed using the Kessler-6 scale (Kessler et al., 2003).

*$p < .10$. **$p < .05$.

1. *Relationship happiness*. This measure is based on a single survey question in response to which each member of the couple rated his or her level of happiness in the relationship on a 0-to-10 scale, with 10 being the happiest. This widely used measure was, as expected, positively correlated with the other relationship quality measures we examined.

2. *Support and affection*. This scale is based on each partner's level of agreement with 12 statements concerning the relationship, such as "My partner shows love and affection to me" and "My partner respects me." The scale ranges from 1 to 4; 4 represents both partners *strongly agreeing* with all 12 statements and 1 represents both partners *strongly disagreeing* with all of them. The measure has strong internal consistency ($\alpha = .94$).

3. *Constructive conflict behavior*. This scale is based on eight survey questions concerning how frequently couples used specific constructive behaviors for managing conflict, such as "Even when arguing, we can keep a sense of humor" and "My partner is good at calming me when I get upset." The scale ranges from 1 to 4; 4 corresponds to *often* exhibiting all the behaviors and 1 corresponds to *never* exhibiting any of the behaviors ($\alpha = .88$).

4. *Avoidance of destructive conflict behavior*. This scale is based on nine survey questions concerning how frequently couples engaged in destructive conflict management behaviors, such as "When we argue, one of us withdraws and refuses to talk about it anymore" and "When we argue, I feel personally attacked by my partner." The scale ranges from 1 to 4; 4 corresponds to *never* exhibiting any of these behaviors and 1 corresponds to *often* exhibiting all these behaviors ($\alpha = .87$).

5. *Fidelity*. This measure is a binary indicator that takes a value of 0 if either member of the couple indicated that he or she or his or her partner had been unfaithful during the follow-up period. Otherwise, it has a value of 1.

We created these five measures by combining the survey responses of both members of the couple into a single couple-level measure. For all the measures except fidelity, we averaged the value of the measure for the mother and the father to create a composite couple-level measure. Partners' assessments of the quality of their relationship were positively correlated; however, reports could differ substantially, with between-partner correlations ranging from .28 to .39 across the four continuous relationship quality measures. We combined mothers' and fathers' reports of relationship quality because we conceptualized these measures as couple level (rather than individual level), with the quality of the couple relationship properly characterized by the average of the quality assessment of the two partners. To confirm the robustness of our results, as part of a sensitivity analysis we examined impacts on mothers' and fathers' reports separately and found very similar results (Moore, Wood, Clarkwest, Killewald, & Monahan, 2012).

With the exception of the fidelity measure, the relationship quality measures were not defined for all sample members. Relationship happiness and the support-and-affection scale were available only for the 59% of couples still romantically involved at the 36-month follow-up. The two conflict behavior scales were available only for the 80% of couples still in regular contact.

Conditioning experimental impact estimates on outcomes that occur after random assignment and that could be influenced by the intervention—such as relationship breakup—has the potential to bias results, because the two research groups may no longer be comparable (McConnell, Stuart, & Devaney, 2008). As described by Moore et al. (2012), we assessed the risk of this sort of truncation bias using established techniques for assessing acceptable levels of sample attrition in experimental evaluations. This assessment indicated that analysis of the two conflict behavior measures has a low risk of bias. Analysis of the relationship happiness and support-and-affection measures had a moderate risk of bias and should thus be interpreted somewhat more cautiously. To address this concern, we also examined an abbreviated six-item version of the support-and-affection scale that could be asked of all couples, including those who were no longer romantically involved. Impact estimates using the full sample and this alternative support-and-affection measure were almost identical to our main results (Moore et al., 2012).

Measures Related to Parenting and Father Involvement

Coparenting relationship. The coparenting measure examined in the 36-month impact

analysis is a single summary index of 10 items drawn from the Parenting Alliance Inventory (Abidin & Brunner, 1995) and was chosen in consultation with one of the developers. These items indicate whether respondents think that they and their partner communicate well in their coparenting roles and are a good coparenting team. Items were asked of all mothers and fathers regardless of whether the couple had remained romantically involved. Using a 5-point scale (ranging from *strongly disagree* to *strongly agree*), respondents were asked to state their level of agreement with the 10 positive statements concerning the coparenting relationship. The scale was created by averaging the responses to the 10 items. The measure has a high level of internal consistency ($\alpha = .95$). The quality of the coparenting relationship used in the impact analysis was defined in a manner parallel to couples' romantic relationship quality, averaging mothers' and fathers' responses to create a couple-level outcome.

Father involvement. Four measures captured fathers' contributions of both time and money to children's well-being:

1. *Father lives with child.* Information from both fathers' and mothers' surveys was used to define a measure of father's residential status at follow-up. The father was considered to live with the focal child (the child born around the time the couple applied for BSF and who made them eligible for the program) only if both parents reported that he did.
2. *Father provides substantial financial support for child.* To measure a father's financial support for children, we relied on mothers' responses to the question "How much of the cost of raising [CHILD] does [FATHER] cover?" The measure is a binary indicator set to 1 if the mother reported that the father covered at least half the cost of raising the child and to 0 otherwise.
3. *Father regularly spends time with child.* Fathers' time with their children was measured on the basis of parents' responses to a question asking how often the father was in contact with the focal child for an hour or more during the previous month. The measure is a binary indicator set equal to 1 if both parents reported that the father spent an hour or more with the focal child every day or almost every day during the previous month and to 0 otherwise.

4. *Father's engagement with child.* Fathers were asked to report the frequency of their engagement with the focal child in 12 activities spanning three domains: (a) caregiving (e.g., feeding or diapering the child), (b) physical play (e.g., rolling a ball or playing chasing games with the child), and (c) cognitive and social play (e.g., singing songs or reading stories). Responses were recorded on a 6-point scale ranging from 6 (*more than once a day*) to 1 (*not at all*). Fathers' responses to the 12 items were averaged to form a single scale of father engagement ($\alpha = .95$).

Measures Related to Child Well-being

Family stability. We examined a binary measure that indicates whether the child had always lived with both parents since birth. The measure is based on both parents' responses.

Economic well-being. We examined three measures in this domain:

1. *Poverty status.* Poverty status is a binary measure indicating whether the monthly income of the child's household—defined as the income of coresidential parents and their coresidential partners, if any—is below the poverty threshold. Respondents were asked about income from earnings, child support, the Supplemental Nutrition Assistance Program (SNAP), Temporary Assistance for Needy Families (TANF), unemployment insurance, disability benefits, money from friends and relatives, and money from any other sources. When children lived with both parents and both parents responded to the survey, the responses of the two parents were averaged when calculating parents' total income.
2. *Material hardship.* As an additional measure of economic distress, we measured material hardship with a binary measure of whether the residential parent reported the child's family experienced any of the following: inability to pay the full amount of rent or the mortgage, having utilities shut off, or eviction. If the child lives with both parents, the family is considered to have experienced material hardship if either parent reports experiencing any of the three hardships.
3. *Reliance on public assistance.* Reliance on public assistance was measured with a binary

variable that indicated whether the child's family received either SNAP or TANF in the past month. If the child lives with both parents, public assistance receipt is indicated if either residential parent reports receiving SNAP or TANF in the past month.

Socioemotional development. We examined BSF's impacts on two measures of children's socioemotional development: (a) behavior problems and (b) socioemotional development. Each of these measures is based on 36-month follow-up survey data collected from the parent who was the primary caretaker of the child (typically the mother).

1. *Behavior problems*. To assess the prevalence of problem behaviors among focal children, the BSF 36-month follow-up survey included 26 items from the version of the Behavior Problems Index (Peterson & Zill, 1986; Zill, 1985) included in the Panel Survey of Income Dynamics, Child Development Supplement. Parents were asked to report whether their children exhibited each of 26 problem behaviors *often*, *sometimes*, or *never*. These items included such behaviors as lying, losing his or her temper easily, demanding a lot of attention, or crying or worrying too much. A summary behavior-problems score was created by averaging responses on all 26 items, with higher scores indicating a greater level of behavior problems ($\alpha = .90$).
2. *Emotional insecurity*. The impact on a child's emotional insecurity in the presence of parental conflict was measured using 10 developmentally appropriate items from the Emotional Reactivity and Behavioral Dysregulation subscales of the Security in the Marital Subsystem—Parent Report Inventory (Davies, Forman, Rasi, & Stevens, 2002). The 80% of parents who at the 3-year follow-up were still in regular contact (defined as seeing or talking to each other at least a few times a month) were asked to rate how frequently various behaviors had occurred during the past month in response to the child seeing arguments or disagreement between the BSF partners, using one of four responses (*often*, *sometimes*, *rarely*, and *never or not applicable*). A summary score was created by averaging responses on all 10 items, with higher scores indicating a greater level of emotional insecurity in the presence of parental conflict ($\alpha = .87$).

Statistical Models

Regression models can improve the statistical precision of impact estimates based on random-assignment research designs and adjust for small initial differences between research groups that can occur by chance or through survey nonresponse. Therefore, all impact estimates presented in this article were generated using regression models that included a large number of baseline control variables, including measures of the couple's initial relationship status and quality, demographic characteristics, and various contextual factors.

These models included data pooled across all eight BSF programs. Each regression model included a series of binary variables indicating each of the eight BSF programs included in the study as well as a set of binary interaction variables indicating whether the couple had applied to a given BSF program and had been assigned to the BSF program group. The program-specific impact estimates are the regression coefficients associated with these interaction variables. The pooled impact estimate for a given outcome was obtained from a mean of the eight program-specific impact estimates in which each program was weighted equally. This method answers the policy-relevant question: What is the average effect of programs that implemented the BSF model? See Moore et al. (2012) for more information on these statistical models.

Estimating Impacts on Couples Who Attended BSF Group Sessions

As described earlier, only 55% of couples assigned to the BSF program group ever attended a group relationship skills session, the program's core service. The primary analysis sample used in this study included all couples who applied for BSF, irrespective of their level of participation in the program. Thus, the impact estimates represent the average effect on all program applicants of being offered BSF services. Such intent-to-treat (ITT) impact estimates are widely used in large-scale random-assignment evaluations for two reasons. First, because everyone randomly assigned is included in the analysis, one can be confident that the program

and control groups were similar at baseline and that statistically significant differences in outcomes that emerge can be attributed to the program. Second, ITT estimates address a policy-relevant research question: What is the effect of offering a program in the "real world," where not everyone in the target population will participate in all program services?

Even so, measuring BSF's impact on couples who actually received the program's core service is also of wide interest. Estimating these treatment-on-the-treated (TOT) effects poses a challenge because couples who attended group sessions were systematically different from those who did not. For example, these couples had higher levels of relationship commitment, were more educated, and were more likely to expect to marry when they applied for BSF (Wood, Moore, & Clarkwest, 2011). Calculating an accurate impact estimate requires that one compare these couples to similar control group couples who would have attended had they been offered access to BSF.

To identify these couples, we used two propensity score matching approaches—a traditional approach and a "likely attender" approach—following recommendations by Schochet and Burghardt (2007) for TOT analysis within an experimental research design. In the traditional approach, program group couples who attended sessions are matched to couples in the control group who appear to have a similar propensity to attend based on a statistical model using their initial characteristics to predict their likelihood of attendance. This method should generate two research groups that are similar in their observed initial characteristics. It is possible, however, that the groups still differ on unmeasured characteristics, such as their level of motivation to improve their relationships. The "likely attender" approach uses propensity scores to identify couples in both research groups who are most likely to attend group sessions if they are offered to them. Thus, these TOT estimates are based not on couples in the BSF program group who actually attended but on couples whose baseline characteristics indicate that they would be likely to attend—regardless of whether they actually attended. This approach ensures that the two research groups are similar on both observed and unobserved characteristics, because both groups are based entirely on initial characteristics and not on post-random-assignment behavior. If the propensity score

model cannot accurately predict who actually attends, however, and the likely attenders in the program group are a substantially different set of couples from those who actually attended, then the results will not yield an accurate estimate of the effects of BSF on those who attended group sessions.

Although these two approaches differ, the reliability of both depends on the extent to which the propensity models identify sample members who would choose to attend group sessions if offered the opportunity. If the predictive power of the model is high, then the two approaches will yield similar results that are likely to reflect BSF's effects on those who attended group sessions. Conversely, if the propensity model has little predictive power, these approaches tend to yield different results, neither of which is likely to represent the program's effects for attenders (Schochet & Burghardt, 2007). Therefore, an examination of the degree to which results from these two methods are similar can suggest how much confidence one can have in the estimates.

For this analysis, we examined impacts on two groups of participants: (a) the 55% of couples who attended at least one group session and (b) the 29% of couples who attended at least half of the group sessions. Examining these two groups allowed us to determine how sensitive BSF impacts appear to be to the size of the "dose" of group sessions. For brevity, we limited our TOT analysis to impacts on the couple relationship. Results were similar for other outcomes.

RESULTS

Impacts on the Couple Relationship

BSF did not make couples more likely to stay together or get married. In fact, BSF couples were somewhat less likely to still be together at the 3-year follow-up, with 57% still romantically involved at this point compared to 60% of control group couples, a difference that is marginally statistically significant (see Table 2). Similarly, BSF couples were somewhat less likely than control group couples to live together (married or unmarried) at the 3-year follow-up (47% and 50%, respectively), a difference that is also marginally statistically significant. BSF and control group couples were equally likely to be married to one another at the 3-year follow-up, with 21% of each group married at this point.

Table 2. *Impacts of the Building Strong Families (BSF) Program on Key Outcomes at the 3-Year Follow–up*

Outcome	BSF group	Control group	Estimated impact	p	Effect size
Relationship status					
Romantically involved (%)	57.4	60.5	−3.2*	.053	−0.079
Living together (married or unmarried) (%)	46.9	49.5	−2.6*	.100	−0.064
Married (%)	20.6	20.9	−0.3	.817	−0.011
Relationship quality					
Relationship happiness scale	8.29	8.30	−0.01	.868	−0.008
Support and affection scale	3.43	3.43	0.00	.989	0.001
Use of constructive conflict behaviors scale	3.22	3.22	−0.01	.770	−0.011
Avoidance of destructive conflict behaviors scale	2.75	2.78	−0.03	.130	−0.054
Neither member of the couple was unfaithful since random assignment (%)	58.2	59.0	−0.8	.628	−0.020
Coparenting					
Quality of coparenting relationship scale	4.19	4.21	−0.02	.510	−0.022
Father's involvement and parenting behavior					
Father lives with child (%)	50.1	51.8	−1.7	.308	−0.040
Father regularly spends time with child (%)	52.4	56.1	−3.6**	.032	−0.089
Mother reports that father provides substantial financial support for raising child (%)	62.8	65.6	−2.8*	.096	−0.072
Engagement with child	4.22	4.26	−0.04	.429	−0.031
Family stability					
Both parents lived with child since birth (%)	42.3	42.7	−0.4	.810	−0.010
Economic well-being					
Family's monthly income below poverty threshold (%)	46.5	46.9	−0.4	.824	−0.010
Family experienced difficulty meeting housing expenses during past year (%)	44.1	44.0	0.1	.956	0.002
Family receiving TANF or SNAP benefits (%)	66.4	65.4	0.9	.564	0.025
Child socioemotional development					
Behavior Problems Index	1.38	1.41	−0.02**	.040	−0.078
Emotional insecurity amid parental conflict	1.40	1.42	−0.02	.430	−0.032
Sample size					
All couples	2,129	2,118			
Couples still romantically involved	1,233	1,253			
Couples still in regular contact	1,713	1,740			
Mothers	1,997	1,984			
Fathers	1,719	1,707			

Note: Data are from BSF 3-year follow-up surveys. The difference between the BSF and control group means may not equal the estimated impact because of rounding. Only the 59% of couples who were still romantically involved were included in the analysis of relationship happiness and support and affection. Only the 79% of couples who were still in regular contact were included in the analysis of conflict management measures. The measure of father provides substantial financial support was based on the mother's report and is defined for couples in which the mother responded to the survey. The measure of father engagement with child was based on the father's report and is defined for couples in which the father responded to the survey. The measure of emotional insecurity amid parental conflict is defined for children of couples who were still in regular contact at the 3-year survey. Other measures are defined for all couples for whom at least one partner responded to the follow-up survey. TANF = Temporary Assistance for Needy Families; SNAP = Supplemental Nutrition Assistance Program.

*p < .10. **p < .05.

BSF did not improve the quality of couples' relationships (see Table 2). Among those who were still romantically involved at the 3-year follow-up, BSF and control group couples reported being equally happy in their romantic relationships, with average ratings of 8.29 and 8.30 respectively on the 0-to-10 relationship happiness scale ($SD = 1.41$ for BSF couples, 1.38 for controls). Similarly, among those still romantically involved, couples in both research groups reported identical levels of supportiveness and affection in their relationships, with average support-and-affection-scale values of 3.43 for both BSF and control group couples on the 1-to-4 scale ($SD = 0.41$ for both groups).

Similarly, BSF did not improve couples' ability to manage their conflicts. Among the 8 in 10 couples who were still in regular contact at the 3-year follow-up, the average score on the 1-to-4 scale measuring the use of constructive conflict behaviors was 3.22 for both BSF and control group couples ($SD = 0.58$ for both groups). There was no difference between the research groups in the avoidance of destructive conflict behaviors among couples who remained in regular contact; the average scale scores were 2.75 for BSF couples and 2.78 for control group couples on the 1-to-4 scale ($SD = 0.61$ for BSF couples, 0.64 for controls). BSF also had no effect on how faithful couples were to each other. At the time of the 36-month follow-up survey, 58% of BSF couples reported no instances of infidelity by either partner since applying for the program, compared with 59% of control group couples, a difference that is not statistically significant.

Impacts on Parenting and Father Involvement

BSF and control group couples reported that their coparenting relationships were of similarly high quality (see Table 2). The average coparenting scale score was 4.19 on the 1-to-5 scale for members of the BSF group and 4.21 for members of the control group, a difference that is not statistically significant ($SD = 0.78$ for BSF couples, 0.77 for controls). These average scores indicate that in both research groups, couples typically agreed or strongly agreed with the 10 positive statements about coparenting included in the scale.

At the 3-year follow-up, BSF had small negative effects on some aspects of father involvement: Fifty-two percent of BSF fathers had spent an hour or more with the focal child on a daily basis during the previous month, compared to 56% of control group fathers, a statistically significant difference. Similarly, 63% of BSF mothers reported that the father covered at least half the cost of raising the child, compared to 66% of mothers in the control group, a marginally statistically significant difference. There was no statistically significant difference between BSF and control group fathers in the likelihood that they lived with their children 3 years after program application (50% and 52%, respectively).

Although BSF fathers spent somewhat less time with their children than fathers in the control group did, BSF had no impact on fathers' self-reported engagement with their children (see Table 2). The average score on the father engagement scale was 4.22 for BSF fathers and 4.26 for control group fathers on the 1-to-6 scale, a difference that is not statistically significant ($SD = 1.25$ for BSF couples, 1.23 for controls). The average scores indicate that, in both research groups, fathers reported that they typically engaged in these activities with their children between a few times a week and once a day.

Impacts on Child Well-being

BSF did not increase the likelihood that children lived with both their biological parents through age 3. At the time of the 3-year follow-up, 42% of BSF children and 43% of children in the control group had lived with both parents continuously since birth (see Table 2). BSF also had no effect on the economic well-being of children. At the 3-year follow-up, 47% of children in both research groups lived in poverty, and 44% of children in both research groups lived in a family that had experienced material hardship during the previous year. In addition, virtually identical percentages of BSF and control group children were living in families that received public assistance (66% and 65%, respectively).

BSF led to modest reductions in behavior problems at the 3-year follow-up. BSF parents reported slightly fewer behavior problems for their children than did parents in the control group. The average Behavior Problems Index score on the 1-to-3 scale was 1.38 for BSF children and 1.41 for children in the control

group ($SD = 0.28$ for both groups), a statistically significant difference that represents an effect size of -0.08. A negative impact on this outcome represents an effect of the program in the desired direction.

Why might BSF have improved children's behavior if it had no positive effects on couples' relationships? One possibility is that this effect originates from services offered by BSF programs that did not specifically aim to improve the couple relationship. Four of the eight BSF programs were also Healthy Families programs. At these four sites, BSF couples received home visits focused on promoting positive parenting behaviors in addition to the relationship skills education offered in BSF group sessions. Control group couples did not receive these home visits. BSF's effect on behavior problems was concentrated in these four Healthy Families programs, suggesting that the effect may be related to the home visits offered in these sites. When impact estimates were pooled across these four programs, the effect size on the Behavior Problem Index was -0.14, which is statistically significant. In contrast, in the four BSF sites that did not offer Healthy Families home visits the pooled effect size was -0.02, not statistically significant. The difference in the impact on this outcome between Healthy Families sites and

other sites is statistically significant (results not shown).

BSF had no effect on children's emotional reactions to parental conflict. Among the 8 in 10 couples in the full research sample who were still in regular contact at the 3-year follow-up, the average score on the composite measure of the child's emotional insecurity amid parental conflict measured on the 1-to-4 scale was 1.40 in the BSF group and 1.42 in the control group ($SD = 0.51$ for both groups). These levels indicate that, in both groups, parents who were still in regular contact typically reported that their children never or only rarely responded to parents' conflicts in these ways.

Impacts for Couples Who Attended Group Sessions

The TOT estimates revealed no strong evidence of effects of BSF on relationship outcomes (either positive or negative) among couples who attended group sessions at all or attended them regularly (see Table 3). Moreover, the two methods we used for estimating these TOT effects—the traditional approach and the "likely attender" method—generally yielded similar results, suggesting that these findings are robust. Among the 55% of couples who attended at least one group session, there were

Table 3. *Sign and Statistical Significance of Impacts on Relationship Outcomes for Couples Attending Building Strong Families Group Sessions*

Outcome	Attended at least one session		Attended at least half of sessions	
	Traditional method	Likely attender method	Traditional method	Likely attender method
Relationship status				
Romantically involved	○	○	○	○
Living together (married or unmarried)	○	○	+	○
Married	○	○	○	○
Relationship quality				
Relationship happiness scale	○	○	○	○
Support and affection scale	○	○	○	○
Use of constructive conflict behavior scale	○	○	○	○
Avoidance of destructive conflict scale	○	○	○	†
Neither member of the couple was unfaithful since random assignment	○	○	○	○

Note: Data are based on information from BSF 3-year follow-up surveys. No impacts were statistically significant at the .05 level. $+$ = statistically significant positive impact at the .10 level; $†$ = statistically significant negative impact at the .10 level; ○ = no statistically significant impact.

no statistically significant effects on any key relationship outcomes using either method. In addition, effect sizes for these estimates were small (the largest was −0.07) and were as likely to be negative as positive (results not shown).

When we estimated impacts for couples who attended at least half of the group sessions, the estimates still showed no impacts of BSF on most relationship quality and status outcomes. However, there are two marginally statistically significant effects—one negative (on avoidance of constructive conflict behaviors, using the "likely attender" method) and one positive (on coresidence, using the traditional matching method)—revealing no clear pattern of impacts. Therefore, BSF did not appear to have a substantial effect on couples' relationships, even among those who attended group sessions regularly. See Moore et al. (2012) for a more detailed discussion of these results.

Separate Impacts for the Eight BSF Programs

By design, the BSF program model gave organizations implementing BSF considerable flexibility in organizing and operating their programs. The goal was to allow for some degree of experimentation with service delivery strategies, because at the time the evaluation began little evidence existed concerning how best to deliver relationship skills education to unmarried parents. As documented by the BSF implementation study, this flexibility led to considerable variation across sites in implementation, with sites choosing different recruitment strategies, different curricula for group sessions, and different approaches to the one-on-one support provided by family coordinators (Dion et al., 2010). Given this program variation across sites, we thought it useful to examine BSF's impacts for the eight programs separately.

At the 15-month follow-up, results varied across the eight sites. The BSF program in Oklahoma City had a consistent pattern of positive effects on relationship outcomes, whereas the Baltimore program had a number of negative effects (Wood, McConnell, Moore, Clarkwest, & Hsueh, 2012). At the 3-year follow-up, the pattern of impacts across sites changed substantially. The negative impacts observed in Baltimore had faded and were generally not statistically significant (results not shown). Similarly, most of the positive effects

in Oklahoma City observed at 15 months did not persist; however, a positive impact on family stability had emerged. At the 3-year follow-up, 49% of BSF children in Oklahoma City had lived with both their biological parents since birth compared to 41% of control group children, a difference that is statistically significant. Although the impacts observed in Baltimore and Oklahoma City faded, numerous negative impacts emerged in the Florida site between the 15-month and 3-year follow-ups. For example, only 55% of BSF couples in Florida were still romantically involved after 3 years, compared with 67% of control group couples. In contrast, at the 15-month follow-up the Florida BSF program had no statistically significant impacts (either positive or negative) on the key outcomes examined. The other evaluation sites generally had little or no effect at either follow-up. Therefore, of the eight evaluation sites, only the program in Oklahoma City showed evidence of positive effects, although these impacts generally did not persist at the 3-year follow-up.

Potential Moderators of Program Effects

Examinations of BSF's average effects might mask heterogeneity in the program's effectiveness for certain subgroups of couples. To evaluate this possibility, we identified key subgroups prior to analysis based on the following baseline characteristics: relationship quality, multiple-partner fertility, fathers' earnings, race and ethnicity, and age. At 3 years, none of these subgroups had a strong pattern of effects (results not shown). The absence of strong subgroup findings at 3 years differs from the pattern found at the 15-month follow-up, when BSF had positive impacts for African Americans on four of the eight primary relationship quality and status measures (Wood et al., 2010). At the 3-year follow-up, BSF did not have a significant impact on any of these eight measures for African American couples. See Moore et al. (2012) for more details on this analysis.

DISCUSSION

BSF represented a new approach to addressing the needs of low-income, unmarried parents and their children. Many new, unmarried parents report that they want and expect to marry each other (Carlson et al., 2005). BSF aimed to

help these parents achieve this goal by offering them services designed to teach relationship skills. The hope was to improve the quality and stability of couples' relationships and ultimately improve outcomes for their children. Although relationship skills education had been shown to be successful in improving relationship quality among middle-class and married couples, the approach had not yet been implemented on a large scale with low-income, unmarried parents, and its effectiveness with this population had not yet been rigorously tested. The BSF evaluation took a thorough look at this approach's potential for success, rigorously testing in multiple locations the effects of relationship skills programs that had been specifically developed for unmarried parents and adapted to their particular needs.

Our results suggest that it is challenging to make this approach work with unmarried parents. At the end of the 3-year follow-up period, BSF had no positive effects on couples' relationships or fathers' involvement with their children. In fact, the program had modest negative effects on some outcomes in these areas. In addition, BSF did not affect most aspects of child well-being that we examined, although it did lead to modest improvements in children's behavior. As discussed earlier, the lack of positive effects on couple outcomes, as well as the concentration of this particular effect in BSF sites that also offered Healthy Families home visits, suggests that this effect was most likely generated not by the group relationship skills education services but by the additional parenting services offered in those sites.

What factors may have limited BSF's success? A key result from the BSF implementation study is that, despite concerted efforts made at the sites, promoting regular attendance at group sessions was a major challenge, and many couples never attended these sessions at all (Dion et al., 2008, 2010). Some have suggested that couples' poor attendance at group sessions was the central reason for the program's lack of positive impacts; however, even among those couples who received a substantial dose of these services we found little evidence of effects on relationship outcomes. These findings suggest that, although regular attendance at group sessions may be a key element of a successful program of this type, it does not guarantee positive impacts. These results also suggest that we need to consider other explanations for BSF's limited success.

The BSF results differ from findings from two other recent studies of similar relationship skills education programs that served low- and moderate-income married couples. As described earlier, a study of a relationship skills program for married military couples, PREP for Strong Bonds, found that the program reduced the likelihood that couples divorced in the 2 years after the program ended (Stanley et al., 2013). In addition, the SHM evaluation, which tested programs similar to BSF but that served low-income married couples, found a pattern of small positive effects on relationship quality but no effect on marriage stability (Hsueh et al., 2012).

Why might relationship skills education programs have less success with unmarried parents than with married couples? One contributing factor may be the relatively low levels of trust and commitment among unmarried low-income parents (Edin & Kefalas, 2005). The behavioral changes required to improve a couple's relationship may involve substantial personal effort. Partners who are less committed to a relationship or distrustful of the commitment of their partner may be more reluctant to do the hard work relationship improvement may require (Van Lange et al., 1997). Thus, on average, unmarried parents may be less likely than married couples to put newly learned relationship skills to use if doing so requires considerable effort on their part and if they are uncertain about their own or their partner's commitment to the relationship. Other differences in the characteristics of married and unmarried parents, such as higher rates of economic disadvantage among unmarried parents and the more frequent occurrence of multiple-partner fertility in these families, may also play a role. These additional stresses may make it difficult for some unmarried parents to focus on putting their newly learned relationship skills to use. Future programs may want to place greater emphasis on directly addressing these stresses.

A noteworthy finding from the BSF evaluation is the fact that a program that aimed to increase relationship stability and father involvement instead led to small reductions in the likelihood that couples remained together and that fathers regularly spent time with their children or provided them with substantial financial support. Perhaps BSF helped some couples with particularly negative or hostile relationships recognize this fact and break up sooner than they otherwise would have, an outcome that may be an appropriate one for these couples.

In addition, qualitative research with BSF couples indicated that the need for fathers to step up and be more responsible was one of the strongest messages that couples took from the program (Dion et al., 2010). This expectation may have led some fathers in particularly disadvantaged circumstances to instead distance themselves from their partner and children. For example, if men do not see themselves as capable of being economically supportive or meeting other expectations of responsible fatherhood, they may reduce engagement with their children in order to protect themselves from a sense of failure or to "shield their children from their own personal failing" (Young, 2011, p. 120). Consistent with that hypothesis, recent research that used BSF data to examine negative impacts of the Baltimore BSF program at 15 months found that BSF fathers at that site were more likely than control group fathers to blame themselves—and especially their own financial, criminal justice, and substance abuse problems—for a relationship breakup, even though their objective outcomes related to earnings, arrests, and substance use were no worse than those of control group fathers (Clarkwest, Killewald, & Wood, in press). Thus, program messages concerning what is involved with being a good father and partner may have led some men to believe they could not meet those expectations and to instead withdraw from these relationships. Future programs serving unmarried parents should give careful attention to the messages they convey to fathers and be sure that goals for good parenting and partnering are presented to fathers in ways that make these goals appear realistic and attainable.

BSF was implemented by eight organizations following a common program model. Seven of the programs did not achieve the central objective of improving couples' relationships. The one exception was the Oklahoma City program, which at the 15-month follow-up had positive effects on relationship quality, romantic involvement, coparenting, and father involvement. These impacts had generally faded by the 3-year follow-up. The Oklahoma program did, however, increase the likelihood that children lived with both their biological parents until age 3. Given that increasing family stability was one of BSF's central goals, this result is noteworthy.

As described earlier, organizations were given flexibility in how they implemented the BSF model, and approaches varied across the eight study sites. Given this fact, as well as the more positive impacts in Oklahoma City, an examination of the distinctive elements of this site's approach is warranted. Oklahoma City was the only BSF program to use the "Becoming Parents" curriculum, which covered similar topics as the other curricula but placed a greater emphasis on the transition to parenthood. The curriculum also prescribed groups twice as large as those recommended in the other two BSF curricula and covered the material in less time (30 rather than 42 hours). The Oklahoma City program offered weekly group sessions in two formats, 3 or 5 hours long, whereas other BSF programs typically offered only 2-hour weekly sessions (Dion et al., 2010). This difference, combined with use of a shorter curriculum, allowed BSF couples in the Oklahoma City program to complete the group sessions in 6 or 10 weeks, whereas couples in other programs needed 5 months to finish. In addition, the Oklahoma City program offered more financial incentives to encourage group attendance than other programs did (Dion et al., 2010). These factors may have played a role in the Oklahoma City site's greater success at getting couples to complete the program. Finally, although only unmarried parents were eligible for the BSF research sample, the Oklahoma City program also served low-income married parents and included both married and unmarried parents in the same group sessions. No other BSF program served parents who were married before their child was conceived. The presence of married couples may have influenced how the group sessions in Oklahoma City functioned as well as how effective they were in improving the outcomes of the couples in the BSF research sample. New programs that plan to offer relationship skills services to unmarried parents may want to examine the approach used in Oklahoma City so that they can build on the program's success. They should also consider adaptations to increase the likelihood that impacts are sustained over the longer term.

The decision to marry can be complex for couples with limited economic prospects. Qualitative research suggests that many low-income couples want both parents to be in a stable economic position before they consider marriage (Cherlin, 2009; Edin & Kefalas, 2005; Edin & Reed, 2005). In addition, recent research on low-income fathers underscores the importance of fathers' perceptions of their economic success in their ability to be engaged and supportive parents

(Young, 2011). These factors may have limited the success of the BSF program model. Recent program efforts have placed greater emphasis on approaches that offer low-income couples both employment and relationship services (Zaveri & Hershey, 2010). ACF is currently sponsoring the Parents and Children Together evaluation, which will examine the effectiveness of programs that offer both employment and relationship services. Perhaps these integrated approaches will have greater success in improving the outcomes of unmarried parents.

REFERENCES

Abidin, R. R., & Brunner, J. F. (1995). Development of a parenting alliance inventory. *Journal of Clinical Child Psychology*, *24*, 31–40. doi:10.1207/s15374424jccp2401_4

Amato, P. (2005). The impact of family formation change on the cognitive, social, and emotional well-being of the next generation. *Future of Children*, *15*, 75–96. doi:10.1353/foc.2005.0012

Carlson, M. J., & Corcoran, M. E. (2001). Family structure and children's behavioral and cognitive outcomes. *Journal of Marriage and Family*, *63*, 779–792. doi:10.1111/j.1741-3737.2001.00779.x

Carlson, M., & Furstenberg, F. F. (2006). The prevalence and correlates of multipartnered fertility among urban U.S. parents. *Journal of Marriage and Family, 68*, 718–732. doi:10.1111/j.1741-3737.2006.00285.x

Carlson, M., McLanahan, S., England, P., & Devaney, B. (2005). *What we know about unmarried parents: Implications for Building Strong Families programs*. Retrieved from http://www.mathematica-mpr.com/publications/PDFs/bsfisbr3.pdf

Center for Research on Child Wellbeing. (2007). *Parents' relationship status five years after a non-marital birth*. Fragile Families Research Brief No. 39, Center for Research on Child Wellbeing, Princeton, NJ.

Cherlin, A. (2009). *The marriage-go-round: The state of marriage and the family in America today*. New York: Knopf.

Clarkwest, A., Killewald, A., & Wood, R. (in press). Stepping up or stepping back: Low income Baltimore parents' responses to the Building Strong Families program. In O. Patterson, & E. Fosse (Eds.), *The cultural matrix: New approaches to the problems of Black youth*. Cambridge, MA: Harvard University Press.

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

Davies, P. T., Forman, E. M., Rasi, J. A., & Stevens, K. I. (2002). Assessing children's emotional security in the interparental relationship. *Child Development*, *73*, 544–562. doi:10.1111/1467-8624.00423

Dion, R. (2005). Healthy marriage programs: Learning what works. *The Future of Children*, *15*, 139–156. doi:10.1353/foc.2005.0016

Dion, R., Avellar, S., & Clary, E. (2010). *The Building Strong Families Project: Implementation of eight programs to strengthen unmarried parent families*. Washington, DC: Mathematica Policy Research.

Dion, R., Hershey, A., Zaveri, H., Avellar, S., Strong, D., Silman, T., & Moore, R. (2008). *Implementation of the Building Strong Families Program*. Washington, DC: Mathematica Policy Research.

Edin, K., England, P., & Linnenberg, K. (2003, September). *Love and distrust among unmarried parents*. Paper presented at the National Poverty Center Conference on Marriage and Family Formation among Low-Income Couples. Retrieved from http://apps.olin.wustl.edu/macarthur/papers/england-LoveDistrust.pdf

Edin, K., & Kefalas, M. (2005). *Promises I can keep: Why poor women put motherhood before marriage*. Berkeley: University of California Press.

Edin, K., & Reed, J. (2005). Why don't they just get married? Barriers to marriage among the disadvantaged. *The Future of Children*, *15*, 117–137. doi:10.1353/foc.2005.0017

Florsheim, P., McArthur, L., Hudak, C., Heavin, S., & Burrow-Sanchez, J. (2011). The Young Parenthood Program: Preventing intimate partner violence between adolescent mothers and young fathers. *Journal of Couple & Relationship Therapy*, *10*, 117–134. doi:10.1080/15332691.2011.562823

Hawkins, A. J., Blanchard, V. L., Baldwin, S. A., & Fawcett, E. B. (2008). Does marriage and relationship education work? A meta-analytic study. *Journal of Consulting and Clinical Psychology*, *76*, 723–734. doi:10.1037/a0012584

Hawkins, A. J., & Fackrell, T. A. (2010). Does relationship and marriage education for lower-income couples work? A meta-analytic study of emerging research. *Journal of Couple & Relationship Therapy*, *9*, 181–191. doi:10.1080/15332691003694927

Hershey, A. A., Devaney, B., Dion, M. R., & McConnell, S. (2004). *Building Strong Families: Guidelines for developing programs*. Princeton, NJ: Mathematica Policy Research.

Hsueh, J., Alderson, D. P., Lundquist, E., Michalopoulos, C., Gubits, D., Fein, D., & Knox, V. (2012). *The Supporting Healthy Marriage Evaluation: Early impacts on low-income families*. New York: MDRC.

Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., & Hiripi, E. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry*, *60*, 184–189.

Kim, H. S. (2011). Consequences of parental divorce for child development. *American Sociological Review*, *76*, 487–511. doi:10.1177/0003122411 407748

Knab, J. T., & McLanahan, S. (2007). Measuring cohabitation: Does how, when, and who you ask matter? In S. L. Hofferth & L. M. Casper (Eds.), *Handbook of measurement issues in family research* (pp. 19–33). Mahwah, NJ: Erlbaum.

Markman, H. J., Renick, M. J., Floyd, F. J., Stanley, S. M., & Clements, M. (1993). Preventing marital distress through communication and conflict management training: A 4- and 5-year follow-up. *Journal of Consulting and Clinical Psychology*, *61*, 70–77. doi:10.1037/0022-006X.61.1.70

Martin, J. A., Hamilton, B. E., Ventura, S. J., Osterman, M. J. K., Wilson, E. C., & Mathews, T. J. (2012). Births: Final data for 2010. *National Vital Statistics Reports, 60*(1). Hyattsville, MD: National Center for Health Statistics.

McConnell, S., Stuart, E. A., & Devaney, B. (2008). The truncation by death problem. *Evaluation Review*, *32*, 157–186. doi:10.1177/019384 1X07309115

McConnell, S., Wood, R., Devaney, B., Avellar, S., Dion, R., & Zaveri, H. (2006). *Building Strong Families: The evaluation design*. Princeton, NJ: Mathematica Policy Research.

McLanahan, S., & Sandefur, G. D. (1994). *Growing up with a single parent: What hurts, what helps*. Cambridge, MA: Harvard University Press.

Moore, Q., Wood, R., Clarkwest, A., Killewald, A., & Monahan, S. (2012). *The long-term effects of Building Strong Families: A relationship skills education program for unmarried parents* (Technical Supplement). Princeton, NJ: Mathematica Policy Research.

Peterson, J. L., & Zill, N. (1986). Marital disruption, parent–child relationships, and behavior problems in children. *Journal of Marriage and the Family*, *48*, 295–307. doi:10.2307/352397

Reardon-Anderson, J., Stagner, M., Macomber, J. E., & Murray, J. (2005). *Systematic review of the impact of marriage and relationship programs*. Washington, DC: Administration for Children and Families, U.S. Department of Health and Human Services.

Schochet, P. Z. (2009). An approach for addressing the multiple testing problem in social policy impact evaluations. *Evaluation Review*, *33*, 539–567. doi:10.1177/0193841X09350590

Schochet, P. Z., & Burghardt, J. (2007). Using propensity scoring to estimate program-related subgroup impacts in experimental program evaluations. *Evaluation Review*, *31*, 95–120. doi:10.1177/0193 841X06288736

Sigle-Rushton, W. & McLanahan, S. (2002). The living arrangements of new unmarried mothers. *Demography*, *39*, 415–433. doi:10.1353/dem. 2002.0032

Stanley, S. M., Rhoades, G. K., Loew, B. A., Allen, E. S., Carter, S., Osborne, L. J., Prentice, D., & Markman, H. J. (2013). *Two-year outcome results from a randomized clinical trial of relationship education with U.S. Army couples*. Manuscript submitted for publication.

Van Lange, P. A. M., Rusbult, C. E., Drigotas, S. M., Arriaga, X. B., Witcher, B. S., & Cox, C. L. (1997). Willingness to sacrifice in close relationships. *Journal of Personality and Social Psychology*, *72*, 1373–1395. doi:10.1037/0022-3514.72.6.1373

Wadsworth, M. E., Santiago, C. D., Einhorn, L., Etter, E. M., Rienks, S., & Markman, H. (2011). Preliminary efficacy of an intervention to reduce psychosocial stress and improve coping in low-income families. *American Journal of Community Psychology*, *48*, 257–271. doi:10.1007/s10464-010-9384-z

Wood, R., McConnell, S., Moore, Q., Clarkwest, A., & Hsueh, J. (2010). *Strengthening unmarried parents' relationships: The early impacts of Building Strong Families*. Princeton, NJ: Mathematica Policy Research.

Wood, R., McConnell, S., Moore, Q., Clarkwest, A., & Hsueh, J. (2012). The effects of Building Strong Families: A healthy marriage and relationship skills education program for unmarried parents. *Journal of Policy Analysis and Management*, *31*, 228–252. doi:10.1002/pam.21608

Wood, R., Moore, Q., & Clarkwest, A. (2011). *BSF's effects on couples who attended group relationship skills sessions: A special analysis of 15-month data*. Princeton, NJ: Mathematica Policy Research.

Young, A. (2011). Comment: Reactions from the perspective of culture and low-income fatherhood. *Annals of the American Academy of Political and Social Science*, *635*, 117–122. doi:10.1177/ 0002716210390316

Zaveri, H., & Hershey, A. (2010). *Initial implementation of a couple-focused employment program*. Washington, DC: Mathematica Policy Research.

Zill, N. (1985). *Behavior problem scales developed from the 1981 Child Health Supplement to the National Health Interview Survey*. Washington, DC: Child Trends, Inc.