Article

Journal of Family Issues
1–29

DOI: 10.1177/0192513X19869385
journals.sagepub.com/home/jfi
SAGE

A Longitudinal Hierarchical Examination of *Smart Steps for Stepfamilies* With Ethnically and Economically Diverse Couples

Katie Reck[1], Brian Higginbotham[2], and Jeffrey Dew[3]

Abstract

Over the past decade, relationship education has grown as a means of enhancing couple relationships. This longitudinal study was designed to examine the experiences and outcomes of 2,828 adults who participated in the *Smart Steps for Stepfamilies: Embrace the Journey* program. Self-report measures of relationship quality, commitment, and relationship instability were administered prior to participation, then immediately, 6 weeks, 6 months, and 1 year after participation. A three-level growth-curve analysis using hierarchical linear model suggested that stepfamily participants experienced small but statistically significant immediate increases in relationship quality and stability, but those gains diminished over time. The statistical analyses tested for moderating effects of individual and couple characteristics on programmatic outcomes over time. Application of these findings and implications for relationship education for stepfamilies are discussed.

[1]Central Michigan University, Mount Pleasant, MI, USA
[2]Utah State University, Logan, UT, USA
[3]Brigham Young University, Provo, UT, USA

Corresponding Author:
Katie Reck, Central Michigan University, 195 Ojibway Court, EHS 412K, Mount Pleasant, MI 48859, USA.
Email: reck1kl@cmich.edu

Keywords
marriage and relationship education, program evaluation, stepfamilies

Scholars and policymakers have called for research investigating the effectiveness of relationship education (RE) to improve the stability and quality of familial relationships (Dion, 2005). Through the healthy marriage initiative (HMI), the U.S. government has dedicated millions of dollars to support healthy marriages and relationships. From this initiative, the past decade has brought a wealth of information related to the efforts contributing to the ongoing debate related to the complex nature of government intervention in effectively supporting healthy family dynamics, particularly among disadvantaged populations (Hawkins, Amato, & Kinghorn, 2013). The publication of evaluative results from these programs generally indicate positive affects among participants, as well as notable nonsignificant findings, indicating some promise as to the effectiveness of such efforts (e.g., Arnold & Beelmann, 2019; Zaveri & Baumgartner, 2016). Notwithstanding, questions remain among practitioners and policymakers as to the long-term success of HMI efforts, particularly among disadvantaged and at-risk populations.

Prior to the first large-scale funding of HMI projects in 2006, RE had primarily targeted European American, middle- and upper-income married couples (Hawkins, Blanchard, Baldwin, & Fawcett, 2008). More recently, evaluative research among high-risk populations (e.g., low-income families) has found small to moderate positive effects, particularly for improving couple communication, relationship quality, and positive coparenting practices (Arnold & Beelmann, 2019; Hawkins & Ooms, 2012). RE specific to stepfamilies is generally less prevalent than programming directed toward premarital or first-marriage couples and empirical evidence documenting the effectiveness of RE with stepfamilies is limited (see Lucier-Greer & Adler-Baeder, 2012).

Stepfamilies are among the fastest-growing family forms in the United States. Estimates from the past decade suggest that between one third and one half of all marriages annually involve previously married individuals (Kreider & Ellis, 2011). The majority of these marriages are likely to involve children from a previous relationship, thereby creating a stepfamily (Whitton, Rhoades, Stanley, & Markman, 2008). Four in 10 adults are estimated to have at least one step relative in their family, either a stepparent, step or half sibling, or stepchild (Pew Research Center, 2011). Stepfamilies in general are at greater risk for relationship instability and poverty compared with married families with only biological children (Adler-Baeder & Higginbotham, 2004; Sweeney, 2010).

These risk factors increase when stepfamily members are also part of other at-risk groups. Over the past two decades, evaluations of federally funded healthy marriage programs have documented positive impacts of RE programs targeted to some at-risk populations, including African American populations (Adler-Baeder, Robertson, & Schramm, 2010). Small yet positive impacts among program participants have also been documented with Latino populations (Higginbotham & Adler-Baeder, 2010). These findings, however, were limited to immediate programmatic outcomes; longitudinal studies are needed to assess the longevity of positive outcomes. Finally, RE programs traditionally targeted middle- to upper-income populations, but now more reach low-income populations (Lundquist et al., 2014). Based on small-to-moderate short-term findings, some scholars posit that low-income couples may similarly improve in relationship outcomes as previously studied moderate-income couples (Hawkins & Fackrell, 2010); however, additional evidence is needed to confirm this supposition. To explore these empirical gaps in the RE literature, we focused on the longitudinal outcomes of a stepfamily RE program that has served thousands of stepfamilies.

Theoretical Perspective

Intervention theory provides a conceptual framework for studying the prevention of human dysfunction. This theory asserts that the negative effects of family dysfunction may be countered by intervention programming that reduces risk factors while strengthening resiliency or protective factors (Coie et al., 1993). This can be done in a variety of ways, including decreasing the problem, interacting with the risk factor to buffer its effects, disrupting the process leading from the risk factor to the problem, and/or by undoing the risk factor. The study of such interventions is important for improving knowledge and best practices.

Intervention theory outlines several themes that are relevant to the present study. First, intervention theory emphasizes the need to incorporate longitudinal designs to assess programming effects over time. Second, individuals are known to adapt to human and environmental interactions; individuals may vary in response to different environments or interventions, which, in turn, may lead to different outcomes. Third, prevention research should incorporate the study of both genders as well as populations that include diverse ethnic and cultural backgrounds. Because dysfunctional behavior may be seen differently based on contextual and cultural factors, interventions may vary in effectiveness based on these factors. Furthermore, Coie et al. (1993) suggest that future research be methodologically rigorous in incorporating adequate sampling, measurements, and appropriate statistical

models, which will ultimately lead to practical implications. These themes provide a rich basis by which interventions can be examined, and like other family stress and resiliency models (e.g., Hill, 1949), they emphasize the fluctuating state of family distress and dysfunction.

Smart Steps for Stepfamilies: Embrace the Journey (hereafter, referred to as *Smart Steps*) is an RE intervention that focuses on risk factors that may influence stepfamily functioning. Within the present study, *Smart Steps* was offered to low-income, ethnically diverse stepfamilies. These populations are at risk for various negative outcomes, including a greater likelihood for relationship dissolution, unrealistic expectations, unclear social norms, and financial difficulties (Ganong & Coleman, 2017). Relationship programming has been theorized to reinforce positive protective factors by instructing participants on these topics and skills (Adler-Baeder & Higginbotham, 2004; Robertson et al., 2006). Drawing on intervention theory, the emphasis of the *Smart Steps* program on stepfamily-specific protective factors should theoretically foster similar positive effects.

Scholars have noted the need to tailor RE program content to the specific needs of the targeted population (Adler-Baeder & Higginbotham, 2004; Larson & Halford, 2011). Accordingly, *Smart Steps* provides a blend of traditional RE content (e.g., conflict resolution and communication skills) as well as stepfamily-specific content (e.g., step- and coparenting skills). The *Smart Steps* program is an evidence-based intervention that was developed using ecological, systems, and family strengths perspectives (Adler-Baeder & Higginbotham, 2004). These theories collectively provide a foundation for conducting intervention programming with individuals in complex family relationships. Ecological systems theories address and integrate the nature of various ecosystems on families (Bronfenbrenner, 1979), thereby placing attention on the immediate micro environments within and across family households as well as on the macro issues of culture and economy. *Smart Steps* addresses micro and macro issues affecting stepfamilies by providing parallel educational programming to both the children and adults in a family unit. The family strengths perspective emphasizes building stepfamily strengths rather than focusing solely on negative dynamics (Gottman & Silver, 1999).

Relationship Education With Stepfamilies

Smart Steps has demonstrated small, yet statistically significant increases in participant-protective factors over the 6-week course and up to 6-weeks after completion of the program; however, no published studies of this program include long-term evaluations of *Smart Steps* participants. Among the

extant studies, *Smarts Steps* has shown short-term increases in stepfamily knowledge, positive communication, agreement on key relational issues (i.e., finances), increased empathy, family engagement, and social support between beginning and completing the program (Higginbotham & Adler-Baeder, 2010; Lucier-Greer, Adler-Baeder, Harcourt, & Gregson, 2014).

Only one meta-analysis of stepfamily RE has been published, and it has reported mostly small but positive effects on family functioning, parenting, and couple relationships (Lucier-Greer & Adler-Baeder, 2012). The majority of the studies predominantly focused on immediate posttest designs. Therefore, the present study assesses long-term outcomes for stepfamily RE and evaluates potential moderators of those outcomes

Based on the extant literature (e.g., Hawkins, Stanley, Blanchard, & Albright, 2012), and the ecological contexts of stepfamily relationships, we account for potential moderators in the present study: age, education, gender, and ethnic background. Age and educational attainment tend to be predictors of marital instability, with those marrying at younger ages and having lower levels of education being at greater risk for divorce (Heaton, 2002). Additionally, men and women may experience RE differently (Schramm & Adler-Baeder, 2012). For example, women with children from a previous marriage are less likely to enter into a new marriage and generally report lower levels of commitment and confidence when they do (Whitton et al., 2008). Finally, among ethnically diverse families, gender differences in familial roles and expectations may affect relationship outcomes (Oropesa & Gorman, 2000). Therefore, moderating factors such as age, education, gender, and ethnic background are important considerations when studying RE program outcomes from both ecological and intervention perspectives.

Study Hypotheses

In light of RE's ability to improve protective factors among other populations (see Hawkins & Ooms, 2012), this study's first hypothesis predicted positive increases in relationship outcomes among stepfamily participants from prior to *Smart Steps* participation (pretest) to 1-year after program participation. These findings would be consistent with previous, nonlongitudinal studies of *Smart Steps* participants as well as with other RE programming demonstrating sustained positive impacts over time (e.g., Halford, Sanders, & Behrens, 2001).

The second hypothesis predicted that individual and couple characteristics would affect postprogram relationship outcomes. Intervention theory allows for differential outcomes (Coie et al., 1993). Likewise, ecological and strength-focused perspectives allow for participant outcomes to differ based

on demographic characteristics. In this study, the individual characteristics of age, gender, number of marriages, ethnicity, and education were examined. Couple characteristics of marital status and household income were also included. Individual and couple effects may play a moderating role in participant experiences with RE, just as relationship outcomes differ by type of relationships (i.e., first marriage vs. remarriage), environmental context (i.e., income), and ethnic background (Hawkins et al., 2013; Rauer et al., 2014).

Method

Data and Sample

Between February 2007 and September 2011, family-service agencies offered 159 *Smart Steps* courses to stepfamilies in a Western state. State and federal grant funding supported the project, which was conducted in accordance with a protocol approved by the research ethics board at a land-grant university. The project included partnerships with 12 family-service agencies located around the state. The partnering agencies recruited stepfamilies via agency listserves, face-to-face invitation, and other marketing strategies. Additionally, the project included statewide advertising efforts in English and Spanish through billboards and public transportation advertisements. In total, 3,186 adults and 2,615 youth (ages 6-17 years) attended a *Smart Steps* class over this 4-year period.

Smart Steps is a 12-hour program that includes six modules of curricular content focusing on the complexities and interdependent nature of stepfamily relationships (Adler-Baeder, 2007). For 6 consecutive weeks, the different agencies held classes for adults and children concurrently for an hour and a half. Then, for approximately 30 minutes of each class, the adult and child groups combined for family strengthening activities. Each session focused on building strengths within the couple and their stepfamily relationships while addressing the unique stressors commonly faced by stepfamilies. Topics included communication, parenting strategies, stress, and financial management. Approximately 6 weeks after completing the *Smart Steps* course, participants returned for a booster class to celebrate the course's completion and to review previously learned course material (see Vaterlaus, Allgood, & Higginbotham, 2012).

Participants completed surveys at five points: the pretest was administered prior to the first class (T1); and posttests were administered immediately following the last class (T2), at the 6-week booster session (T3), at 6 months after program completion (T4), and again at 1 year after program completion (T5). For T1 to T3, participants completed surveys via paper and pencil at the intervention administration site. However, starting with T4, the survey team

Table 1. Smart Steps Participants Demographic Characteristics: Individual and Couple Variables.

	Women			Men		
Individual variables	*M*/%	*SD*	Range	*M*/%	*SD*	Range
Age	32.62	7.263	17-66	34.93	8.064	18-65
Education	12.98	2.769	0-34	12.894	2.820	0-29
No. of marriages	1.48	0.859	0-6	1.43	0.895	0-7
Ethnicity						
African American	1			1.1		
Asian American	0.6			0.5		
Caucasian	66.6			66.9		
Hispanic/Latino	27.4			27.6		
Native American	1.3			0.8		
Biracial	0.2			1.3		
Unknown	1.6			0.2		
Other	1.3			1.7		
Marital status						
Married	62.6			63.6		
In an unmarried relationship	37.4			36.4		
Household income ($)						
<28,500	37.6			34.1		
28,500-50,000	27.4			28.2		
50,001-72,500	16.0			18.3		
72,501-109,000	13.4			13.5		
>109,001	5.5			5.8		

mailed surveys to participants. Unfortunately, this change in survey modality resulted in a large attrition rate after T3. We address this issue in the Results and Study Limitations sections.

Of the 3,186 adult participants who participated, not all consented to be part of the surveys and 156 self-reported as single parents. The sample used in this study include the 2,828 adults who provided data and self-identified as living in a stepfamily, meaning they or their partner was currently parenting at least one child from a previous relationship. Approximately two thirds of adults reported being in a married relationship; the other third reported being in an unmarried, cohabitating relationship (see Table 1). A majority of participants in this study included couple data ($n = 1,317$ couples), meaning both partners attended the *Smart Steps* course and completed study surveys individually. Approximately 13% of study participants were individuals both

coupled and living in a stepfamily but whose partner did not attend the course or complete survey data. The sample was composed of 1,316 men (46.5%) and 1,512 women (53.5%). Survey response rates were 88% at T2, 91% at T3, 23% at T4, and 27% at T5. We chose to model the growth curve using a three-level hierarchical linear model (HLM). We chose this analysis for this study, in part, due to its ability to correctly model individual and couple data even when participants do not complete all survey waves. For a description of how missing data were handled in HLM, see Raudenbush and Bryk (2002).

Participants reported their demographic characteristics at T1. Women ranged in age from 17 to 66 years ($M = 32.6$, $SD = 7.3$), reported between 0 and 34 years of formal education ($M = 13.0$, $SD = 2.8$), and had been married 0 to 6 times ($M = 1.5$, $SD = 0.9$; 62.6% were married at the time of participation). The majority of women self-identified as either Caucasian (62.6%) or Hispanic/Latina (27.4%). More than a third (37.6%) of women reported an annual household income below $28,500, and another 27.4% reported $28,500 to $50,000; only 5.5% reported annual household incomes above $109,000. Men ranged in age from 18 to 65 years ($M = 34.9$, $SD = 8.1$), reported between 0 and 29 years of formal education ($M = 12.9$, $SD = 2.8$), and had been married 0 to 7 times ($M = 1.4$, $SD = 0.9$; 63.6% were married at the time of participation). Most men self-identified as either Caucasian (66.9%) or Hispanic/Latino (27.6%). More than a third (34.1%) of men reported an annual household income below $28,500, and another 28.2% reported $28,500 to $50,000; only 5.8% reported annual household incomes above $109,000. In general, participants were in their 30s, had a high school education, and had two children living in the household. A majority were in their first or second marriage, averaging 4 years in length; unmarried couples were together on average for 2 to 3 years. Despite the state's relatively homogeneous ethnic and racial demographics, more than one third reported not being Caucasian. The sample also included a diverse range of household incomes with more than half making less than $50,000 per year.

Measures

Relationship Constructs. We used three relationship outcome measures to assess change in couples' relationships over time—specifically, relationship quality, couple commitment, and relationship instability. These measures were determined based on the interests of the funders and targeted outcomes of the *Smart Steps* program.

Relationship quality. Relationship quality was assessed using an adapted version of the Quality Marriage Index, which is a five-item scale adapted

from Norton (1983). On a 7-point Likert-type scale, participant responses ranged from *very strongly disagree* to *very strongly agree* in answer to five statements: (1) "We have a good relationship"; (2) "My relationship with my partner is very stable"; (3) "Our relationship is strong"; (4) "My relationship with my partner makes me happy"; and (5) "I really feel like part of a team with my partner." The Likert-type response options for each statement range from *very strongly disagree* (scored as 1) to *very strongly agree* (7). Overall scores are calculated by summing the responses for a total possible range of 5 to 35; higher scores correspond with higher relationship quality. Internal reliability (Cronbach's alpha) was high for each of the five data collection points ($.96 < \alpha < .98$).

We tested for measurement invariance in relationship quality across time. The men's relationship quality measures met the criteria for strong measurement invariance (i.e., similar factor loadings and intercepts over time; Widaman, Ferrer, & Conger, 2010). Women's relationship quality measures met the standard for weak measurement invariance (i.e., similar factor loadings over time but intercepts that did not remain stable over time).

Relationship commitment. We measured commitment to the relationship using a subscale of the Commitment Inventory (Stanley & Markman, 1992), this scale comprises four statements: (1) "My relationship with my partner/spouse is more important to me than almost anything else in my life," (2) "I may not want to be with my partner/spouse a few years from now" (reversed coded), (3) "I like to think of my partner/spouse and me more in terms of 'us' and 'we' than 'me' and 'him/her,'" and (4) "I want this relationship to stay strong no matter what rough times we may encounter." Five-point Likert-type response options for each statement range from *strongly disagree* (1) to *strongly agree* (5). Scores were calculated by summing the responses for a total possible range of 4 to 20; higher scores correspond with higher relationship commitment. Cronbach's alpha was less stable across the five data collection points than for the other two relationship constructs ($.68 < \alpha < .82$). We tested for measurement invariance in commitment across time. Neither women's commitment nor men's commitment met the Widaman et al.'s (2010) criteria for weak invariance. Because of our hypotheses, we still tested commitment, but we are mindful of this major limitation on this particular dependent variable.

Relationship instability. Relationship instability was assessed using the total score of a scale developed by Booth, Johnson, and Edwards (1983), consisting of four questions (1) "Have you ever thought your relationship might be in trouble?" (2) "Has the thought of getting a divorce or separation crossed your

mind?" (3) "Have you discussed divorce or separation with a close friend?" and (4) "Have you or your partner/spouse ever seriously suggested the idea of divorce or separation?" Three-point Likert-type response options for each statement range from *never* (1) to *yes, recently* (3). Overall scores were calculated by summing the responses for a total possible range of 4 to 12; higher scores correspond with higher relationship instability. Cronbach's alpha was high for each of the five data collection points ($.88 < \alpha < .89$). We tested for measurement invariance in relationship instability across time. Both women's instability and men's instability met the Widaman et al.'s (2010) criteria for strong invariance.

Time. To examine the longitudinal effects of the outcome variables, we measured time based on the *Smart Steps* course and survey schedule (i.e., T1-T5). To most accurately portray the elapsed time between the T2 and T5 data points, we created time values using the T2 survey as a baseline and giving each data collection point a time value represented by fractions of a year labeled (i.e., T1 = −0.12, T2 = 0, T3 = 0.12, T4 = 0.5, and T5 = 1). Furthermore, three variables were created to account for the possible shape of the association between time and the outcome variables: linear (time), quadratic ($time^2$), and cubic time ($time^3$). This allowed for the examination of linear, curvilinear, and a second possible fluctuation in outcome variables over time. The time variable was mean centered prior to creating the quadratic and cubic terms so that they would not be collinear with each other.

Individual and Couple Characteristics. We gathered individual and couple demographic characteristics with the presurvey. Individual demographic data included age, gender, number of marriages, ethnicity, and education. Gender was dummy coded with men as the reference group. Number of marriages was dummy coded into three variables to distinguish groups among categorical variables, resulting in four subsequent variables: never married, second marriage, and higher order marriage (3+ marriages), with first marriage used as the reference group. The ethnicity variable was similarly dummy coded into Latino and "other" ethnicity variables with European American set as the reference group. Finally, a "number of surveys completed" variable was created to sum the total surveys completed by each individual to capture differences in participant retention over time.

Couple-level demographic data, collected at the presurvey, included marital status and household income. Marital status distinguished whether an individual was in a married or unmarried relationship (i.e., cohabitating). This variable was dummy coded with those in a married relationship being the reference group. The second couple-level variable, household income,

was calculated based on the combined responses from two survey questions: "Approximately what is your personal income per year?" and "Approximately what is your partner's or spouse's total personal income per year?" Based on these questions, five household income categories were created: (1) $0 to $28,500, (2) $28,500 to $50,000, (3) $50,001 to $72,500, (4) $72,501 to $109,000, and (5) households making more than $109,001.

Analyses

We modeled the growth curve of the three relationship outcomes using a three-level, multimodel hierarchical linear approach. HLM, also known as multilinear or multilevel modeling, allows data with a hierarchical structure to be analyzed by nesting. That is, first-level units are nested within second-level and third-level units (e.g., students nested within classrooms, and classrooms nested within schools). This type of analysis allows researchers to assess how different layers of data interact and are associated with dependent or outcome variables (Raudenbush & Bryk, 2002).

HLM accounts for the variance in the dependent variable at the lowest level (e.g., specific time points), while considering information from all other levels (e.g., the individual and then individuals within couples; Raudenbush & Bryk, 2002). In HLM, the lowest level of individual variables predicts the dependent variables. The upper-level variables then predict the intercept and slopes of the lower-level equations.

We chose to model participants' growth curves in HLM because the sample violates several assumptions required for other types of longitudinal analyses. First, participant observations and error structures within the present study are not independent of one another (i.e., the same individual has multiple data points over time, and the individuals are linked as couples). Second, the time points were not uniformly spaced, and some longitudinal techniques (e.g., repeated measures analysis of variance) assume that the time points have a uniform spacing. Third, not every participant completed every wave of data collection; we had a high level of attrition, especially following T3. With the exception of the high levels of attrition, HLM accommodates all of these issues and produces unbiased estimates (see Raudenbush & Bryk, 2002). We conducted some follow-up analyses with a subsample of participants who completed T4 and/or T5 to ensure that we could replicate our findings from the overall sample (see the Results section and online Appendix 1 for more information).

For the first level of analysis, the linear, quadratic, and cubic measures of time were examined. Level-1 variables were group-mean centered (i.e., centered with the group at each time point) to improve the interpretation and

variance of the intercept as well as to mitigate potential coefficient bias (Raudenbush & Bryk, 2002). The intercept and linear time coefficient were free to vary. The other two Level-1 effects were fixed. The equation used for the Level-1 model for time was

$$\text{Outcome variable}_{tij} = \pi_{0ij} + \pi_{1ij} * (\text{Time}) + \pi_{2ij} * (\text{Time}^2) + \pi_{3ij} * (\text{Time}^3) + e_{tij}.$$

The second level of analyses included nesting Level-1 (time) variables within individual characteristics to identify any moderating effects on the outcome variables. That is, individual-level (i.e., Level 2) variables were used to model the Level-1 intercept and slopes. Predicting the intercept based on individual-level variables is important because it identifies the individual characteristic variables that associated with outcome differences present at T2 (baseline). The Level-2 individual characteristics consisted of variables unique to each participant (i.e., age, gender, number of marriages, ethnicity, education, and number of surveys completed). Level-2 variables were group-mean centered with fixed effects. The full Level-2 model equations were

$$\begin{aligned}\pi_{0ij} = {} & \beta_{00j} + \beta_{01j} * (\text{age}_{ij}) + \beta_{02j} * (\text{gender}_{ij}) + \beta_{03j} * (\text{never married}_{ij}) \\ & + \beta_{04j} * (\text{second marriage}_{ij}) + \beta_{05j} * (\text{higher order marriage}_{ij}) \\ & + \beta_{06j} * (\text{Latino}_{ij}) + \beta_{07j} * (\text{other ethnicity}_{ij}) + \beta_{08j} * (\text{education}_{ij}) \\ & + \beta_{09j} * (\text{surveys completed}_{ij}) + r_{0ij}\end{aligned}$$

$$\begin{aligned}\pi_{1ij} = {} & \beta_{10j} + \beta_{11j} * (\text{age}_{ij}) + \beta_{12j} * (\text{gender}_{ij}) + \beta_{13j} * (\text{never married}_{ij}) \\ & + \beta_{14j} * (\text{second marriage}_{ij}) + \beta_{15j} * (\text{higher order marriage}_{ij}) \\ & + \beta_{16j} * (\text{Latino}_{ij}) + \beta_{17j} * (\text{other ethnicity}_{ij}) \\ & + \beta_{18j} * (\text{education}_{ij}) + \beta_{19j} * (\text{surveys completed}_{ij})\end{aligned}$$

$$\begin{aligned}\pi_{2ij} = {} & \beta_{20j} + \beta_{21j} * (\text{age}_{ij}) + \beta_{22j} * (\text{gender}_{ij}) + \beta_{23j} * (\text{never married}_{ij}) \\ & + \beta_{24j} * (\text{second marriage}_{ij}) + \beta_{25j} * (\text{higher order marriage}_{ij}) \\ & + \beta_{26j} * (\text{Latino}_{ij}) + \beta_{27j} * (\text{other ethnicity}_{ij}) \\ & + \beta_{28j} * (\text{education}_{ij}) + \beta_{29j} * (\text{surveys completed}_{ij})\end{aligned}$$

$$\pi_{3ij} = \beta_{30j} + \beta_{31j} * (\text{age}_{ij}) + \beta_{32j} * (\text{gender}_{ij}) + \beta_{33j} * (\text{never married}_{ij}) + \beta_{34j} * (\text{second marriage}_{ij}) + \beta_{35j} * (\text{higher order marriage}_{ij}) + \beta_{36j} * (\text{Latino}_{ij}) + \beta_{37j} * (\text{other ethnicity}_{ij}) + \beta_{38j} * (\text{education}_{ij}) + \beta_{39j} * (\text{surveys completed}_{ij})$$

The final level of analysis included time and individual characteristics nested within couple characteristics (i.e., marital status and annual household income) to identify any moderating effects on outcome variables. These Level-3 couple variables included marital status and couple household income. The main intercept was allowed to vary; the rest of the effects were fixed. The third-level equations were

$$\beta_{00j} = \gamma_{000} + \gamma_{001} (\text{marital status}_{j}) + \gamma_{002} (\text{household income}_{j}) + u_{00j}$$

$$\beta_{10j} = \gamma_{0100} + \gamma_{0101} (\text{marital status}_{j}) + \gamma_{0102} (\text{household income}_{j})$$

$$\beta_{20j} = \gamma_{200} + \gamma_{201} (\text{marital status}_{j}) + \gamma_{202} (\text{household income}_{j})$$

$$\beta_{30j} = \gamma_{300} + \gamma_{301} (\text{marital status}_{j}) + \gamma_{302} (\text{household income}_{j})$$

Results

Relationship Quality

The raw means of relationship quality among participants showed a slight increase in relationship quality from T1 to T3 ($M_{\text{change}} = 0.44$), then declines from T3 to T5 ($M_{\text{change}} = -0.15$). Table 1 provides the HLM output of the relationship quality model.

The predicted intercept coefficient for relationship quality was 5.52 ($p < .001$). The only Level-2 variable that statistically enhanced the prediction of relationship quality was gender ($B = -0.10$, $p < .05$); as a whole, men had a 0.10 higher score than women on relationship quality at T2. Among Level-3 variables, only household income statistically enhanced the prediction of relationship quality at T2 ($B = 0.08$, $p < .05$); couples with higher levels of income tended to have slightly higher relationship quality scores than those with lower income levels (see Figure 1).

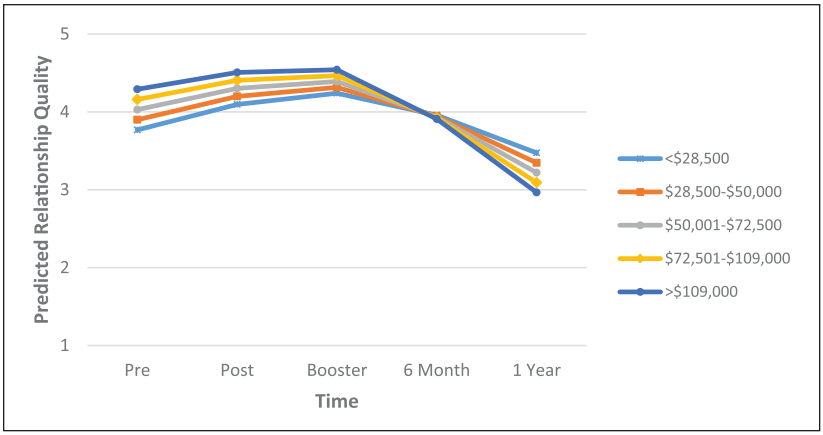

Figure 1. Predicted relationship quality: Household income interaction.

The linear ($B = 1.02$, $p = .001$), quadratic ($B = -4.95$, $p = .001$), and cubic ($B = 3.49$, $p = .001$) time slopes were each statistically significant, suggesting that, similar to the raw mean score data, the HLM model follows a curvilinear trend over time. The positive change from T1 to T3 represents a quarter of a point change in standard deviation increase, but these gains attenuated back to baseline levels over time.

Household income was the only Level-3 variable that statistically enhanced prediction of the linear time slope for relationship quality ($B = -0.23$, $p < .001$). Graphing the predicted values (see Figure 1) showed that people with lower household income levels were predicted to have steeper increases in relationship quality over time. These individuals also had less steep decreases in relationship quality over time compared with those at higher income levels. Those with the highest levels of income reported the highest levels of relationship quality at T2, but they experienced the greatest decrease in predicted relationship quality scores a year later (T5).

Relationship Commitment

Raw mean scores of commitment depicted an increase in couple commitment over time, but only slightly; commitment increased 0.08 from T1 to T3, followed by a decline of 0.01 at T4 and an increase of 0.01 at T5. Table 2 provides the results of the HLM output. No statistically significant slope changes in couple commitment was found over time.

Table 2. Hierarchical Model Predicting Relationship Quality, Couple Commitment, and Relationship Instability Over Time.

	Relationship quality				Couple commitment				Relationship instability			
Fixed effects	*B*	*SE*	*t*-ratio	*df*	*B*	*SE*	*t*-ratio	*df*	*B*	*SE*	*t*-ratio	*df*
Intercept (π_0)	5.52***	0.07	75.88	1,187	4.33***	0.03	129.38	1,298	1.79***	0.04	51.14	1,298
Marital status (γ_{001})	−0.02	0.06	−0.27	1,187	−0.15***	0.03	−4.82	1,298	0.02	0.03	0.54	1,298
Household income (γ_{002})	0.08**	0.02	3.15	1,187	0.06***	0.01	5.60	1,298	−0.04***	0.01	−3.24	1,298
Age (γ_{010})	0.00	0.01	−0.77	778	0.00	0.00	−1.18	835	0.00	0.00	−1.33	835
Gender (γ_{020})	−0.10**	0.04	−2.76	778	−0.03	0.02	−1.28	835	0.07***	0.02	4.40	835
Never married (γ_{030})	0.07	0.09	0.83	778	0.07	0.05	1.37	835	−0.01	0.04	0.18	835
Second marriage(γ_{040})	0.03	0.06	0.54	778	0.03	0.03	0.90	835	−0.01	0.03	−0.42	835
Higher order marriage (γ_{050})	0.07	0.10	0.66	778	0.02	0.06	0.33	835	0.03	0.04	0.77	835
Latino (γ_{060})	0.15	0.12	1.33	778	0.01	0.07	0.07	835	−0.04	0.05	−0.76	835
Other race (γ_{070})	0.03	0.13	0.23	778	−0.10	0.08	−1.25	835	−0.05	0.06	−0.91	835
Education (γ_{080})	0.00	0.01	−0.37	778	0.00	0.01	0.51	835	0.01**	0.01	2.60	835
Surveys completed (γ_{0100})	−0.05	0.04	−1.20	778	−0.05	0.03	−1.84	835	0.05**	0.02	2.57	835
Linear slope (π_1)	1.02***	0.20	5.18	1,670	0.19	0.11	1.71	1,828	−0.23*	0.09	−2.44	1,828
Marital status (γ_{101})	−0.25	0.18	−1.37	1,670	−0.16	0.10	−1.60	1,828	0.31***	0.09	3.56	1,828
Household income (γ_{102})	−0.23***	0.06	−3.66	1,670	−0.05	0.04	−1.46	1,828	0.02	0.03	0.73	1,828
Age (γ_{110})	0.03	0.03	0.86	1,670	0.00	0.02	0.17	1,828	0.00	0.02	0.08	1,828
Gender (γ_{120})	0.21	0.22	0.94	1670	−0.03	0.12	−0.27	1,828	−0.02	0.10	0.23	1,828
Never married (γ_{130})	0.10	0.55	0.18	1,670	−0.39	0.30	−1.30	1,828	−0.37	0.26	−1.43	1,828
Second marriage(γ_{140})	0.37	0.30	1.22	1,670	0.22	0.17	1.28	1,828	−0.01	0.14	−0.05	1,828
Higher order marriage (γ_{150})	−0.19	0.56	−0.34	1,670	−0.21	0.31	−0.67	1,828	0.44	0.26	1.70	1,828
Latino (γ_{160})	0.71	0.69	1.03	1,670	0.54	0.38	1.41	1,828	−0.44	0.32	−1.34	1,828
Other race (γ_{170})	0.19	0.79	0.24	1,670	−0.49	0.41	−1.18	1,828	0.35	0.35	1.00	1,828
Education (γ_{180})	−0.07	0.07	−1.07	1,670	−0.02	0.04	−0.44	1,828	−0.02	0.03	−0.46	1,828
Surveys completed (γ_{1100})	−0.43*	0.22	−1.96	1,670	0.11	0.13	0.85	1,828	0.07	0.11	0.68	1,828

(continued)

Table 2. (continued)

Fixed effects	Relationship quality				Couple commitment				Relationship instability			
	B	*SE*	*t*-ratio	*df*	*B*	*SE*	*t*-ratio	*df*	*B*	*SE*	*t*-ratio	*df*
Quadratic slope (π_2)	−4.95***	0.96	−5.15	1,670	−0.54	0.55	−0.99	1,828	0.23	0.46	0.52	1,828
Marital status (γ_{201})	0.83	0.88	0.94	1,670	−0.22	0.50	−0.44	1,828	−1.46***	0.42	−3.43	1,828
Household income (γ_{202})	0.42	0.32	1.31	1,670	−0.07	0.19	−0.39	1,828	0.15	0.16	0.95	1,828
Age (γ_{210})	0.09	0.16	0.57	1,670	0.00	0.09	0.00	1,828	−0.05	0.08	−0.67	1,828
Gender (γ_{220})	−1.08	1.05	−1.03	1,670	−0.92	0.60	−1.53	1,828	−0.15	0.51	−0.29	1,828
Never married (γ_{230})	1.87	2.53	0.74	1,670	0.37	1.47	0.25	1,828	−0.52	1.24	−0.42	1,828
Second marriage (γ_{240})	0.84	1.55	0.54	1,670	−0.86	0.88	−0.98	1,828	0.01	0.74	−0.01	1,828
Higher order marriage (γ_{250})	−2.67	2.73	−0.98	1,670	−1.98	1.51	−1.31	1,828	−0.16	1.27	−0.13	1,828
Latino (γ_{260})	−1.89	3.21	−0.59	1,670	0.20	1.84	0.11	1,828	1.62	1.55	1.05	1,828
Other race (γ_{270})	−3.71	3.81	−0.98	1,670	−1.15	2.13	−0.54	1,828	3.02	1.79	1.68	1,828
Education (γ_{280})	−0.22	0.33	−0.65	1,670	−0.02	0.19	−0.11	1,828	−0.02	0.16	−0.14	1,828
Surveys completed (γ_{2100})	0.42	1.14	0.37	1,670	−0.17	0.67	−0.25	1,828	0.46	0.56	0.83	1,828
Cubic slope (π_3)	3.49***	1.05	3.33	1,670	0.20	0.57	0.35	1,828	0.24	0.48	0.49	1,828
Marital status (γ_{301})	−0.34	0.97	−0.35	1,670	0.70	0.54	1.31	1,828	0.99*	0.45	2.20	1,828
Household income (γ_{302})	−0.10	0.35	−0.38	1,670	0.15	0.19	0.78	1,828	−0.24	0.16	−1.48	1,828
Age (γ_{310})	−0.16	0.17	−0.89	1,670	0.02	0.09	−0.17	1,828	0.05	0.08	0.68	1,828
Gender (γ_{320})	0.81	1.14	−0.71	1,670	1.05	0.63	1.66	1,828	−0.26	0.53	−0.48	1,828
Never married (γ_{330})	−1.79	2.85	−0.63	1,670	0.24	1.58	0.15	1,828	1.00	1.33	0.75	1,828
Second marriage (γ_{340})	−0.90	1.61	−0.56	1,670	0.71	0.88	0.81	1,828	−0.11	0.74	−0.14	1,828
Higher order marriage (γ_{350})	2.72	2.92	0.93	1,670	2.01	1.57	1.28	1,828	−0.21	1.32	−0.16	1,828
Latino (γ_{360})	1.13	3.50	0.32	1,670	−0.81	1.93	−0.42	1,828	−0.97	1.62	−0.60	1,828
Other race (γ_{370})	4.63	4.15	1.18	1,670	2.05	2.23	0.92	1,828	−3.62*	1.88	−1.93	1,828
Education (γ_{380})	0.35	0.36	0.97	1,670	0.10	0.20	0.51	1,828	0.02	0.17	0.09	1,828
Surveys completed (γ_{3100})	0.00	1.19	0.00	1,670	−0.15	0.67	−0.23	1,828	−0.55	0.56	−0.97	1,828

Note. *SE* = standard error; *df* = degrees of freedom.
*$p < .05$. **$p < .01$. ***$p < .001$.

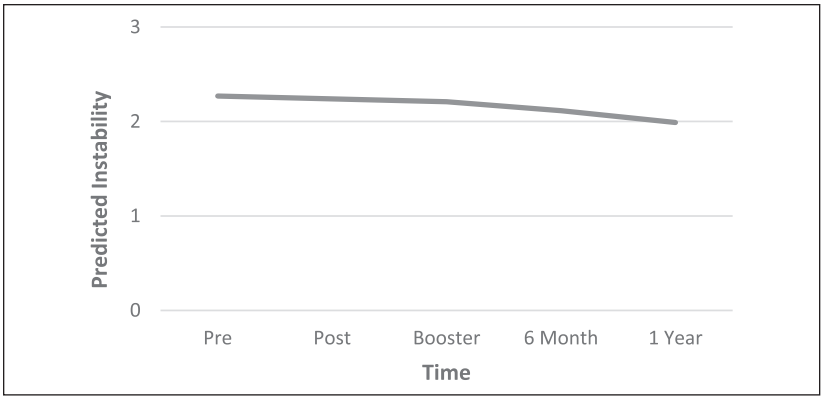

Figure 2. Predicted relationship instability: Time effect.

Relationship Instability

Relationship instability scores showed similar results as the couple commitment outcome variable, with some improvement as instability scores decreased slightly over time ($M_{change} = 0.10$; see Figure 2). The HLM output is presented in Table 2, showing a statistically significant intercept ($B = 1.77$, $p < .001$) as well as linear time effects ($B = -0.22$, $p < .05$). Quadratic and cubic changes over time were not statistically significant. Thus, over time, mean relationship instability scores decreased linearly by a mean of 0.22.

At Level 2, gender ($B = 0.07$, $p < .001$), education ($B = 0.01$, $p < .05$), and the number of surveys completed ($B = 0.05$, $p < .05$) were associated with the intercept (i.e., statistically enhanced the prediction of instability). At Level 3, the only variable that statistically enhanced the prediction of instability was household income ($B = -0.04$, $p < .001$). This means that men, individuals with higher education levels, those who completed more program surveys, and couples with higher income levels had lower levels of instability at postprogram.

No Level-2 individual characteristic variables statistically enhanced prediction of instability in the linear model, meaning all groups had similar linear trajectories over time. At Level 3, however, marital status statistically enhanced prediction of instability in the linear ($B = -0.31$, $p < .01$), quadratic ($B = -1.46$, $p < .01$), and cubic time terms ($B = 0.99$, $p < .05$). Predicted levels of instability showed that cohabiting couples started lower, peaked at about the same point as married couples, then declined to lower levels of instability over time (see Figure 3). This suggests that cohabiting

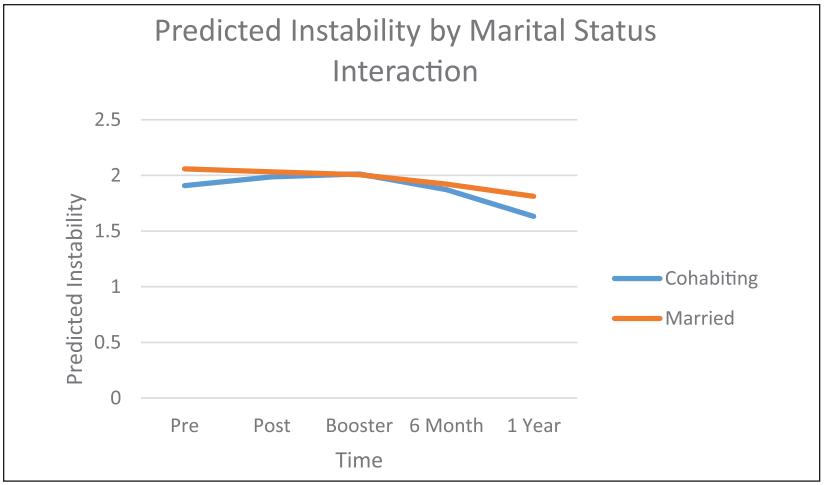

Figure 3. Predicted relationship instability: Marital status interaction.

couples benefitted more from the intervention vis-á-vis relationship instability than did married couples.

Follow-Up Analyses

The survey mode change caused our attrition rates to be high by T4 and T5 points. We reasoned that this issue may have influenced the trends found in this study. Consequently, we conducted some exploratory post hoc analyses (see Appendix 1, available online, for detailed description and results tables) to examine the effect of the mode change and attrition on participant outcomes.

One important finding from our post hoc analysis showed that the likelihood of completing T4 and/or T5 was positively associated with T1 relationship quality. Given these findings, we attempted to replicate and further explore this finding but only with participants who completed T4 and/or T5 surveys. Thus, we created a subsample of those participants who provided data in these later waves ($n = 319$ women and 231 men).

To briefly summarize our post hoc findings, we were able to replicate the relationship quality slopes in the subsample, but only for those participants who had characteristics that made them most likely to drop out of the study, yet these participants still completed T4 and/or T5 surveys anyway. That is, for those who were most likely to drop out based on the logistic regression analysis, but in fact did not, we found that their relationship quality increased

up to T3, then declined. This decline, however, did not reach to the average baseline, suggesting that these participants experienced a net gain as a result of participating in the program. By way of contrast, those who were most likely to complete T4 and/or T5 surveys, or who had middling probabilities of staying in, evidenced no real change in slope. We stress that these findings are exploratory and not definitive (e.g., we may have statistical power issues that limit our ability to find smaller effects in the data) but raise interesting questions as to who may benefit most from RE. Again, for a more full summation of the follow-up analysis and graphs, see Appendix 1 (available online).

Discussion

This study provides insights into the effectiveness of a stepfamily RE program that serves large numbers of ethnically and economically diverse participants. These results build on mixed findings from the general RE literature showing both positive and nonsignificant changes in participant outcomes (e.g., Arnold & Beelmann, 2019; Hawkins & Fackrell, 2010; Zaveri & Baumgartner, 2016). Consistent with our first hypothesis, results showed that participant relationship quality increased over time; however, this increase was small with an effect size of only a quarter of a point change in standard deviation. These changes also attenuated by 1-year postprogram. Follow-up analysis further suggests that these positive results may be heavily influenced by more robust positive changes experienced by a subsample of participants who were predicted to drop out of the study but did not. For example, those who had the least stable relationships at T1 were predicted to have the highest dropout rates; however, among those in these relationships who remained in the intervention, they experienced notable positive changes over time unlike those with more stable relationships at T1 (see Table A2 in Appendix 1, available online). These findings may suggest that RE may be most beneficial among those who most need the intervention, yet retaining these individuals throughout the entirety of the program remains a significant barrier.

Overall, modeled results in this study showed modest improvement in relationship quality scores for all participants from pretest to the booster session 6 weeks after program completion. This finding is consistent with other more rigorous comparison study models of the *Smart Steps* program (Lucier-Greer et al., 2014). However, after the booster session in this study, participants showed declines to near pretest scores 1 year after completing the program. These results once again are consistent with previously published meta-analyses showing short-term, postassessment relationship improvements (Hawkins et al., 2008; Hawkins & Fackrell, 2010; Lucier-Greer &

Adler-Baeder, 2012) as well as reported long-term attenuation of relationship quality outcomes, particularly for high-risk samples (Hawkins et al., 2008).

Although statistically significant time effects were found for relationship quality and instability from pre- to postprogram, it is important to note that these effects were not maintained a year after participation (T5), except for those who were predicted to not complete all waves of data in the follow-up analysis. The findings from this study raises questions about long-term program efficacy. One suggestion to improve long-term outcomes is to increase intervention dosage. In this *Smart Steps* program, participants could receive up to 13 hours (six 2-hour session plus 1-hour booster session) of in-person RE, a moderate dosage. Among RE that targets low-risk groups, between 9 and 20 dosage hours are suggested to ensure programmatic efficacy (Hawkins et al., 2012). The short-term positive outcomes we observed suggest additional dosage, and follow-up programming may be necessary to ensure lasting program efficacy. *Smart Steps* is currently a 12-hour program, so a moderate increase in intervention dosage is reasonable and may help foster long-term improvement. That said, further research is needed to identify whether and how optimal dosage hours vary between low-risk and high-risk populations.

RE practitioners and scholars have emphasized the utility of expanding follow-up RE to ensure sustained program effects (Markman & Rhoades, 2012). RE should not be viewed as a one-time treatment service, rather, as a means to ensure the continued practice of gained knowledge and skills. Follow-up programming may be particularly important for disadvantaged populations who are at elevated risk for experiencing multiple stressors, such as poverty, relationship dissolution, multiple transitions, and poor child and family outcomes (e.g., Ganong & Coleman, 2017). This may be especially true for the stepfamilies and ethnically diverse couples who were the target of this *Smart Steps* program. The positive, short-term effects found in this study parallel other *Smart Steps* findings showing short-term improvements, particularly among less stable couples (Lucier-Greer et al., 2014). Providing ongoing follow-up programming after the 6-week booster may help sustain positive program effects by further reinforcing learned skills and knowledge (see Vaterlous, Allgood, & Higginbothm, 2012). One strategy as suggested by Cordova et al. (2014) eliminates the single postprogram booster session strategy. Instead, they suggest implementing follow-up RE as one would an annual health "check-up"—that is, providing regular RE dosages in smaller, more flexible intervals to ensure consistent long-term results. The difficulty in implementing this strategy, however, lies in retaining couples for multiple booster classes and funding for these services.

Our second hypothesis was that relationship outcomes vary over time according to individual and couple characteristics. This hypothesis was based on existing literature, for example, that low-income couples are at greater risk than their more financially stable counterparts to experience divorce, marital instability, marital discord, and dissatisfaction (Lichter & Carmalt, 2009). This hypothesis was supported in two cases: household income for relationship quality and marital status for relationship instability. Although couples with the highest incomes tended to report the highest levels of relationship quality to begin, by 1 year after the program, these couples tended to have the lowest relationship quality. Couples with lower household income experienced a more sustained program effect and, notably, had experienced the most growth and the least decline of any other socioeconomic group a year after program participation. Generally, we would expect those with higher incomes to report higher quality relationships, therefore, it is somewhat surprising to see such a decline in what "should" be relatively stable relationships. This finding provides new evidence specific to stepfamilies that is consistent with previous findings among more general audiences, indicating that RE may work best with higher risk populations (Amato, 2014), including ethnically diverse, low-income stepfamilies. This study adds to the literature suggesting that lower income couples in stepfamilies may benefit more from RE over time than couples with fewer marital risk factors, which is consistent with intervention theories theme of addressing or removing risk factors while increasing resiliency and strengths (Coie et al., 1993). Importantly too, improvement in relationship quality of at-risk, low-income couples is a goal consistent with a family strengths perspectives and the purposes of the HMI (e.g., Hawkins & Sage, 2015).

Some scholars have suggested that the decline in relationship quality among RE participants (e.g., higher income couples in this study) may be the result of newly raised standards and expectations of couple relationships. This new "standard" as it were of what a healthy relationship "should be," which may have had no impact on the couple prior to the intervention, may prove difficult for couples to live up to (Schramm, Galovan, & Goddard, 2017). Although not directly tested in this study, this concept may be true for some couples in this study. However, intervention theory asserts that individuals should be expected to vary in response to interventions; therefore, it is important to examine subgroup and long-term outcomes to determine differences in intervention outcomes. Even among a homogeneous group, such as stepfamilies in this study, evaluative outcomes may differ when more closely examined. In the future, practitioners should be cognizant of the possibility of RE introducing new stressors to couples and should ensure that

program content adequately addresses such stressors. Calls by Schramm et al. (2017) for a more balanced approach between skills-based and principle-based programming may be an appropriate strategy for such cases.

For the couple commitment outcome, no statistically significant time effects were found in this study indicating little to no change in predicted participation scores over time. For relationship instability, a statistically significant linear effect was found indicating participants' instability scores declined over time. Some studies have shown statistically significant positive effects for stability (e.g., Stanley et al., 2014); however, the positive, albeit small effects of the current study add to the promise of current RE efforts to improve relationship stability among diverse family groups (e.g., ethnic minority, stepfamilies).

A second time interaction in this study was found between the linear time effect of relationship instability and marital status. In Figure 3, married couples experienced a more consistent effect in terms of instability decline. Cohabiting couples though experienced a greater overall benefit from the intervention than married couples in terms of relationship instability. Given the increasing trend among couples to cohabit prior to marriage, or choosing not to marry at all (Isen & Stevenson, 2010), this finding is promising since RE will likely continue to serve this growing stepfamily population. Less is known about cohabitating stepfamilies than married stepfamilies—an area in need of future exploration (Ganong & Coleman, 2017).

For stepfamily RE, a reexamination of curricular content may enlighten scholars and practitioners as to what is most beneficial to this growing population. For example, traditional RE topics of healthy communication and conflict management as well as stepfamily-specific content including stepparenting may be advisable; however, other topics such as clarifying parenting roles, setting familial expectations, and commitment may be equally important to long-term cohabitating couples' stability (Markman & Rhoades, 2012).

In addition to the time interaction effects in this study, statistically significant differences between gender and household income were found at T1 on relationship quality and instability variables. Finding gender differences in RE outcomes is not uncommon (e.g., Amato, 2014); however, the majority of RE literature has not found gender differences in program outcomes over time (e.g., Hawkins et al., 2008). Our findings suggest that although men and women in stepfamilies may differ in reported relationship quality and stability prior to program participation, they can experience similarly positive short- and long-term trajectories during and following participation in RE programs.

Finding statistically significant gender differences in this study may further be explained by the shared identity of being in a step relationship. This

shared experience may limit the prevalence of other group differences. For example, low- and high-income participants exhibit group relationship trajectory differences over time within RE (Hawkins & Fackrell, 2010); however, couples in stepfamilies may connect and empathize on common issues unique to stepfamilies, thereby tempering some common gender differences. Similarly, this study's findings suggest that individuals from diverse stepfamily backgrounds may experience comparable relationship outcomes in RE.

Relationship quality and instability also significantly differed according to household income prior to program participation. Low socioeconomic status is known to be associated with poor family outcomes (e.g., Lichter & Carmalt, 2009). Similarly, T1 differences in education for relationship instability were not surprising given the literature on early marriage, out-of-wedlock childbearing, and relationship dissolution (Fein, Burstein, Fein, & Lindberg, 2003).

Finally, the number of surveys completed was a statistically significant moderator for relationship instability, meaning those who completed fewer surveys reported more relationship instability. This is important in considering a possible selection effect that can potentially explain some of our findings. For example, those with less stable relationships may have been more prone to dissolve those relationships and consequently prematurely ceased study participation, possibly changing residences during the postsurvey time period and, therefore, could not be located. This could skew our findings to the extent that RE outcomes vary according to relationship quality. More generally with regard to retention, the more demanding the RE program participation is in terms of time commitment, the greater the likelihood of participant dropout or refusal to attend the program at all (Busby, Larson, Holman, & Halford, 2015).

Study Limitations

Particular strengths of the present study include the large and diverse participant sample, and the longitudinal examination of a stepfamily RE program serving economically and ethnically diverse families. However, our results should be interpreted with due caution given, among other limitations, the lack of a control group and the sample attrition over time. Notably, the attrition pattern suggests that a great deal of the attrition was due to a change of survey collection procedure from in-person to mail starting at T4.

Other study limitations are noteworthy. First, although European American and Latinos were well represented in the sample, other ethnic groups were underrepresented, and additional research is needed to examine the extent to which these findings apply to other racial and

ethnic classifications. Second, possible ceiling and selection effects may be prevalent in the present study given the tendency of participants to report relatively high commitment and low instability. This suggests that, although at-risk populations were targeted during sample recruitment, the sample obtained may have skewed toward those with healthier relationships. Third, this sample was limited to those who reported being in a current stepfamily relationship but included those who did and did not participate in the *Smart Steps* course together with that partner. It is unclear whether there was a cumulative effect; that is, whether the number of partners participating (one or both) moderated outcomes. Finally, the parental status of participating individuals (i.e., biological parent, stepparent, or both) was unknown, and this too may have moderated outcomes.

Additional limitations exist within the methods. First, some of the selected measures had participants agree with statements that focused on the couple relationship "ever" being in trouble. Future studies may consider measurement constructs that specifically focus on current relationship needs or difficulties. Second, although the response rate for survey responses were high at post- and 6-week booster surveys, a significant decline was experienced in 6-month and 1-year mail follow-ups. This decline is likely reflective in the change of survey collection procedure from in-person to mail, as well as the mobility of high-risk populations. Exact dosage of RE for participants was not included in this study (e.g., did they arrive late or leave early), neither was whether couples attended classes together or separately. These variables may affect overall experience in RE courses and reported outcomes and should be included in future study. Finally, the combined household income variable was created based on the combination of partners' self-reported earnings. This may not accurately reflect the total household income of a couple (e.g., child support).

Implications for Research and Practice

Several implications for stepfamily RE can be derived from this study. First, although the overall effect of program participation on relationship quality was relatively small, it was large enough to have practical importance. For example, Wolf (1986) suggested that an effect size of .25 represents a practical difference among education programs. Thus, the long-term postprogram impact observed in this study can inform discussions regarding the value of RE and its ability to sustain healthy relationships among at-risk populations (Hawkins et al., 2013; Johnson, 2013). Furthermore, the exploratory follow-up analysis conducted in this study suggests that these small effect changes may in fact be due to larger changes seen within subgroups or samples within

the data (e.g., those predicted to drop out of the study, but who did not). Future studies should focus on specific subgroups within study samples to determine what key characteristics lend to improved/maintained program outcomes. Second, careful consideration regarding program dosage is needed. This program provided a moderate 13-hour dosage for participants, which has historically resulted in positive program outcomes among traditional RE populations (Hawkins et al., 2012). With stepfamily couples who are at greater risk for negative family outcomes, a higher RE dosage and expanded follow-up programming may be needed to ensure sustained improvements.

This study's findings also suggest a need to refine RE recruitment and retention methods to reach distressed high-risk couples. Although ethnically diverse and low-income stepfamilies were targeted for inclusion in the program, preparticipation mean scores indicate that a more resilient and potentially less distressed sample was obtained than expected. This selection effect is common within RE programming (cf. Stanley, 2001). Strategic community partnerships or mandated courses (e.g., several states have divorce education requirements) may be the best ways to reach distressed individuals and couples for RE. In terms of retention, there should be considerations for both the in-person program time as well as the follow-up research components of RE projects. In an effort to reduce barriers to ongoing participation in the current study, specific efforts were taken to help families attend in-person programming. Specifically, participation was free of charge, available in multiple locations across the state, and included dinner for the entire family prior to each class. Consequently, retention was acceptable from T1 to T3. The problem with attrition occurred after the in-person programming was complete. Additional efforts and supports are necessary to retain participants in subsequent evaluations.

A particularly important finding was that relationship trajectories over time only varied according to household income and marital status. Historically, RE was implemented among mostly homogeneous middle-class, European American populations (Dion, 2005; Ooms & Wilson, 2004). Findings from this study suggest that *Smart Steps* is able to serve a more heterogeneous population of stepfamilies (e.g., of various ethnicities and marital statuses), which is pertinent when considering future curriculum development and implementation. The implication is that practitioners of stepfamily RE can effectively implement programming with diverse stepfamily populations.

Conclusion

This study adds to existing RE literature by examining the long-term effects of couples in a stepfamily RE program. Relationship quality and stability

improved among participants, particularly in the short term, but effects attenuated over time; household income and marital status also moderated program effects over time. Future studies of stepfamily RE may build on these findings by further examining how initial positive improvements in RE can be sustained.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Funding for this research was provided in part by the U.S. Department of Health and Human Services, Administration for Children and Families (Grant No. 90FE0129 and Grant No. 90YD0227). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the U.S. Department of Health and Human Services, Administration for Children and Families.

Supplemental Material

Supplemental material for this article is available online.

References

Adler-Baeder, F., & Higginbotham, B. J. (2004). Implications of remarriage and stepfamily formation for marriage education. *Family Relations*, *53*, 448-458. doi:10.1111/j.0197- 6664.2004.00053.x

Adler-Baeder, F. (2007). *Smart Steps: Embrace the journey*. Retrieved from http://www.stepfamilies.info/smart-steps.php

Adler-Baeder, F., Robertson, A., & Schramm, D. G. (2010). Community education programs serving couples in stepfamilies: A qualitative study of format, content, and service delivery. *Journal of Extension*, *48*(5), 1-7. Retrieved from https://www.joe.org/joe/2010october/rb1.php

Amato, P. R. (2014). Does social and economic disadvantage moderate the effects of relationship education on unwed couples? An analysis of data from the 15-month building strong families evaluation. *Family Relations*, *63*, 343-355. doi:10.1111/fare.12069

Arnold, L. S., & Beelmann, A. (2019). The effects of relationship education in low-income couples: A meta-analysis of randomized-controlled evaluation studies. *Family Relations*, *68*, 22-38. doi:10.1111/fare.12325

Booth, A., Johnson, D., & Edwards, J. N. (1983). Measuring marital instability. *Journal of Marriage and the Family*, *45*, 387-394. doi:10.2307/351516

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.

Busby, D., Larson, J. H., Holman, T. B., & Halford, W. K. (2015). Flexible delivery approaches to couple relationship education: Predictors of initial engagement and retention of couples. *Journal of Child and Family Studies*, *24*, 3018-3029. doi:10.1007/s10826-014-0105-3

Coie, J. D., Watt, N., West, S. G., Hawkins, D., Asarnow, J., Markman, H., . . . Long, B. (1993). The science of prevention: A conceptual framework for and some directions for a national research program. *The American Psychologist*, *48*, 1013-1022. doi:10.1037//0003-066X.48.10.1013

Cordova, J. V., Eubanks Fleming, C. J., Ippolito, M. M., Hawrilenko, M., Sollenberger, J. W., Harp, A. G., & Wachs, K. (2014). The marriage checkup: A randomized controlled trial of annual relationship health checkups. *Journal of Consulting and Clinical Psychology*, *82*, 592-604.

Dion, M. R. (2005). Healthy marriage programs: Learning what works. *Marriage and Child Wellbeing*, *15*, 139-156. doi:10.1353/foc.2005.0016

Fein, D. J., Burstein, N. R., Fein, G. G., & Lindberg, L. D. (2003). *The determinants of marriage and cohabitation among disadvantaged Americans: Research findings and needs*. Cambridge, MA: Abt Associates.

Ganong, L. H., & Coleman, M. (2017). *Stepfamily relationship: Development, dynamics and interventions*. New York, NY: Springer.

Gottman, J. M., & Silver, N. (1999). *The seven principles for making marriage work*. New York, NY: Three Rivers Press.

Halford, W. K., Sanders, M. R., & Behrens, B. C. (2001). Can skills training prevent relationship problems in at-risk couples? Four-year effects of a behavioral relationship education program. *Journal of Family Psychology*, *15*, 750-768. doi:10.1037/0893-3200.15.4.750

Hawkins, A. J., Amato, P. R., & Kinghorn, A. (2013). Are government-supported healthy marriage initiatives affecting family demographics? A state-level analysis. *Family Relations*, *62*, 501-513. doi:10.1111/fare.12009

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

Hawkins, A. J., & Ooms, T. (2012). Can marriage and relationship education be an effective policy tool to help low-income couples form and sustain healthy marriages and relationships? A review of lessons learned. *Marriage and Family Review*, *48*, 524-554. doi:10.1080/01494929.2012.677751

Hawkins, A. J., Stanley, S. M., Blanchard, V. L., & Albright, M. (2012). Exploring programmatic moderators of the effectiveness of marriage and relationship education programs: A meta-analytic study. *Behavior Therapy*, *43*, 77-87. doi:10.1016/j.beth.2010.12.006

Heaton, T. B. (2002). Factors contributing to increasing marital stability in the United States. *Journal of Family Issues*, *23*, 392-409. doi:10.1177/0192513X02023003004

Higginbotham, B. J., & Adler-Baeder, F. (2010). Enhancing knowledge and agreement among ethnically and economically diverse couples in stepfamilies with the Smart Steps: Embrace the Journey program. *Journal of Extension*, *48*, 1-5.

Hill, R. (1949). *Families under stress: Adjustment to the crises of war separation and reunion*. New York, NY: Harper.

Isen, A., & Stevenson, B. (2010). *Women's education and family behavior: Trends in marriage, divorce and fertility*. Cambridge, MA: National Bureau of Economic Research.

Johnson, M. D. (2013). Optimistic or quixotic? More data on marriage and relationship education programs for lower income couples. *American Psychologist*, *68*, 111-112. doi:10.1037/a0031793

Kreider, R. M., & Ellis, R. (2011). *Number, timing, and duration of marriages and divorces: 2009* (Current Population Reports, P70-125). Retrieved from http://www.census.gov/prod/2011pubs/p70-125.pdf

Larson, J., & Halford, W. K. (2011). One size does not fit all: Customizing couple relationship education for unique couple needs. In J. L. Wetchler (Ed.), *Handbook of clinical issues in couple therapy* (2nd ed., pp. 293-309). New York, NY: Routledge.

Lichter, D. T., & Carmalt, J. H. (2009). Religion and marital quality among low-income couples. *Social Science Research*, *38*, 168-187. doi:10.1016/j.ssresearch.2008.07.003

Lucier-Greer, M., & Adler-Baeder, F. (2012). Does couple and relationship education work for individuals in stepfamilies? A meta-analytic study. *Family Relations*, *61*, 756-769. doi:10.1111/j.1741-3729.2012.00728.x

Lucier-Greer, M., Adler-Baeder, F., Harcourt, K. T., & Gregson, K. D. (2014). Relationship education for stepcouples reporting relationship instability—evaluation of the Smart Steps: Embrace the Journey curriculum. *Journal of Marital and Family Therapy*, *40*, 454-469. doi:10.1111/jmft.12069

Lundquist, E., Hsueh, J., Lowenstein, A. E., Faucetta, K., Gubits, D., Michalopoulos, C., & Knox, V. (2014). *A family-strengthening program for low-income families: Final impacts from the supporting healthy marriage evaluation*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Markman, H. J., & Rhoades, G. K. (2012). Relationship education research: Current status and future directions. *Journal of Marital Family Therapy*, *38*, 169-200. doi:10.111/j.1752-0606.2011.00247.x

Norton, R. (1983). Measuring marital quality: A critical look at the dependent variable. *Journal of Marriage and the Family*, *45*, 141-151. doi:10.2307/351302

Ooms, T., & Wilson, P. (2004). The challenges of offering relationship and marriage education to low-income populations. *Family Relations*, *53*, 440-447. doi:10.1111/j.0197-6664.2004.00052.x

Oropesa, R. S., & Gorman, B. K. (2000). Ethnicity, immigration, and beliefs about marriage as a "Tie that binds." In L. J. Waite & C. Bachrach (Eds.), *The ties that bind: Perspectives on marriage and cohabitation* (pp. 188-211). Hawthorne, NY: Walter de Gruyter.

Pew Research Center. (2011). *A portrait of stepfamilies*. Retrieved from http://www.pewsocialtrends.org/2011/01/13/a-portrait-of-stepfamilies/

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models* (2nd ed.). Thousand Oaks, CA: Sage.

Rauer, A. J., Adler-Baeder, F., Lucier-Greer, M., Skuban, E., Ketring, S. A., & Smith, T. (2014). Exploring processes of change in couple relationship education: Predictors of change in relationship quality. *Journal of Family Psychology*, *28*, 65-76. doi:10.1037/a0035502

Robertson, A., Adler-Baeder, F., Collins, A., DeMarco, D., Fein, D., & Schramm, D. (2006). *Meeting the needs of married, low-income stepfamily couples in marriage education services*. Cambridge, MA: Abt Associates.

Schramm, D. G., & Adler-Baeder, F. (2012). Marital quality for men and women in stepfamilies: Examining the role of economic pressure, common stressors, and stepfamily-specific stressors. *Journal of Family Issues*, *33*, 1373-1397. doi:10.1177/0192513X11428126

Schramm, D. G., Galovan, A. M., & Goddard, H. W. (2017). What relationship researchers and relationship practitioners wished the other knew: Integrating discovery and practice in couple relationships. *Family Relations*, *66*, 696-711. doi:10.1111/fare.12270

Stanley, S. M. (2001). Making the case for premarital education. *Family Relations*, *50*, 272-280. doi:10.1111/j.1741-3729.2001.00272.x

Stanley, S. M., & Markman, H. J. (1992). Assessing commitment in personal relationships. *Journal of Marriage and the Family*, *54*, 595-608. doi:10.2307/353245

Stanley, S. M., Rhoades, G. K., Loew, B. A., Allen, E. S., Carter, S., Osborne, L. J., Prentice, D., & Markman, H. J. (2014). A randomized controlled trial of relationship education in the U.S. Army: 2-year outcomes. *Family Relations*, *63*, 482-495. doi:10.1111/fare.12083

Sweeney, M. (2010). Remarriage and stepfamilies: Strategic sites for family scholarship in the 21st century. *Journal of Marriage and Family*, *72*, 667-684. doi:10.1111/j.1741-3737.2010.00724.x

Vaterlaus, J. M., Allgood, S. M., & Higginbotham, B. J. (2012). Stepfamily Education Booster Sessions. *Social Work with Groups*, *35*, 150-163. doi:10.1080/01609513.2011.599017

Whitton, S. W., Rhoades, G. K., Stanley, S. M., & Markman, H. J. (2008). Effects of parental divorce on marital commitment and confidence. *Journal of Family Psychology*, *22*, 789-793. doi:10.1037/a0012800

Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives*, *4*, 10-18. doi:10.1111/j. 1750-8606.2009.00110.x

Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Beverly Hills, CA: Sage.

Zaveri, H., & Baumgartner, S. (2016). *Parents and children together: Design and implementation of healthy marriage programs* (OPRE Report Number 2016-63). Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.