Impact Evaluation of

U.S. Committee for Refugees and Immigrants' (USCRI) Refugee Family Strengthening (RFS) Program in 11 Different Program Sites in the U.S.

Final Impact Evaluation Report for:

U.S. Committee for Refugees and Immigrants

September 28, 2020

Prepared by: Educational Evaluators, Inc.

Jennifer Salerno, DNP

Sarah Rockhill, Aida Hernandez, MA, LPC

Michael Young, PhD

Assumptions, Constraint, & definitions

Recommended Citation:

Salerno, J. Rockhill, S. Hernandez, A. Young, M. (2020). *Impact Evaluation of* U.S. Committee for Refugees and Immigrants' (USCRI) Refugee Family Strengthening (RFS) Program in 11 Different Program Sites in the U.S. Educational Evaluators, Inc.

Acknowledgements:

The RFS Program was a collaborative project developed and administered with the assistance of Program Officers and Program Coordinators of the U.S. Committee for Refugees and Immigrants and evaluated by Educational Evaluators, Inc.

Disclosure:

Jennifer Salerno, Sarah Rockhill, Aida Hernandez, and Michael Young declare they have no conflicts of interest.

This publication was prepared under Grant Number 90FM0092 from the Office of Family Assistance (OFA) within the Administration for Children and Families (ACF), U. S. Department of Health & Human Services (HHS). The views expressed in this report are those of the authors and do not necessarily represent the views or policies of HHS, ACF, or OFA.

Structured Abstract: "The Evaluation of U.S. Committee for Refugees and Immigrants' (USCRI) Refugee Family Strengthening (RFS) Program in 11 Different Program Sites in the U.S."

Objective. To examine the impact of the Refugee Family Strengthening Program (RFS). The study examined the effects of two interventions, based on RFS, on improvements in three outcome measures: Outcome 1: Marriage and relationship skills, Outcome 2: Conflict management skills, and Outcome 3: Economic stability.

Study design. The study used a randomized control trial (RCT) design. Participants self-identified as a refugee or immigrant, and enrolled as individuals, couples, or families. Participants were randomized in a 1:1:1 ratio to one of two intervention groups or a control group. In no case was one member of a couple, or family unit assigned to one group and other member(s) assigned to a different group. Participants in both intervention groups received 16 hours of instruction, plus case management. The instruction consisted of Relationship Enhancement (both groups), Family Stress and Conflict Management (Intervention I), or Financial Management (Intervention II). Controls received case management only. The testing instrument was a self-report questionnaire that included scales measuring communication/relationship skills, conflict management skills, and economic stability. All participants, regardless of their assignment, completed a pretest questionnaire, and a follow-up questionnaire three months after program completion, about six months after enrollment. Intervention participants completed a post-test questionnaire soon after they completed the program. Data were analyzed using analysis of covariance. Separate analyses were conducted by participant type. Covariates included in the models were the pretest score for the outcome variable under consideration and demographic characteristics that failed tests of baseline equivalence.

Results. At six-month follow-up *individual* participants in both Intervention I and Intervention II and *couple* participants in Intervention I scored significantly better than controls on communication/relationship skills, conflict management skills, and economic stability. *Couple* participants in Intervention II scored better than controls for the marriage and relationship skills outcome and the conflict management skills outcome, but no better than controls for the economic stability outcome. *Family* participants in both Intervention I and Intervention II did not score better than controls for any of the outcome measures.

Conclusion. Evaluations of programming for refugees and immigrants that include interventions promoting relationship enhancement have been sorely lacking. The RFS program provides comprehensive healthy relationship programming based on skill-building and fostering family communication and bonding. Thus, this study, which documents the positive effects of the RFS programming with refugees, plays something of a pioneering role and should be of value to other agencies that work with these populations.

Contents

I. INTRODUCTION....1
A. Introduction and study overview....1
B. Primary research question(s)....5
II. INTERVENTION AND COUNTERFACTUAL CONDITIONS....7
A. Description of program as intended....7
B. Description of counterfactual condition as intended....10
C. Research Questions about the intervention and counterfactual conditions as implemented....11
III. STUDY DESIGN....13
A. Sample formation and research design....13
B. Data collection....14
IV. ANALYSIS METHODS....18
A. Analytic sample....18
B. Outcome Measures....24
C. Baseline equivalence and sample characteristics....26
V. FINDINGS AND ESTIMATION APPROACH....37
A. Implementation evaluation....37
B. Primary impact evaluation....39
C. Sensitivity analyses....45
D. Additional analyses....47
VI. DISCUSSION....49
REFERENCES....53
VIII. APPENDICES....56
A. Logic Model (or Theory of Change) for the program....56
B. Data and Study Sample....56
C. Attrition Rates and Baseline Equivalence of the RCT Design....56
D. Additional Analyses....56
Appendix A: Logic Model (or Theory of Change) for the program....57
Appendix B: Data, Sample, and Measures....58
Appendix C: Randomization assignment by site and participant type....62
Appendix D: Data preparation....64

Tables

II.1. Description of intended intervention and counterfactual components and target populations 9

II.2. Staff training and development to support intervention and counterfactual components 10

II.3. Research questions of the implementation study, by implementation element 11

IV.1a. Individual sample sizes by intervention status: Intervention I and Control Group Participants 18

IV.1b. Individual sample sizes by intervention status: Intervention II and Control Group Participants 19

IV.1c. Cluster and individual sample sizes by intervention status: Couple Participants, Treatment I, and Control Group 20

IV.1d. Cluster and individual sample sizes by intervention status: Couple Participants, Treatment II and Control Group 21

IV.1e. Cluster and individual sample sizes by intervention status: Family Participants: Intervention I and Control Group 22

IV.1f. Cluster and individual sample sizes by intervention status: Family Participants: Intervention II and Control Group 22

IV.1g. Rates of attrition for Intervention I compared to Control by outcome 23

IV.1h. Rates of attrition for Intervention II compared to Control by outcome 24

IV.2. Outcome measures used for primary impact analyses research questions 25

IV.3a. Summary statistics of key baseline measures and baseline equivalence across study groups, for individual participants completing six-month follow-up questionnaire: Intervention I and Control Group 27

IV.3b. Summary statistics of key baseline measures and baseline equivalence across study groups, for INDIVIDUAL participants completing six-month follow-up questionnaire: Intervention II and Control Group 28

IV.3c. Summary statistics of key baseline measures and baseline equivalence across study groups, for COUPLE participants completing six-month follow-up questionnaire: Intervention I and Control Group 30

IV.3d. Summary statistics of key baseline measures and baseline equivalence across study groups, for COUPLE participants completing six-month follow-up questionnaire: Intervention II and Control Group 32

IV.3e. Summary statistics of key baseline measures and baseline equivalence across study groups, for FAMILY participants completing six-month follow-up questionnaire: Intervention I and Control Group 33

IV.3f. Summary statistics of key baseline measures and baseline equivalence across study groups, for FAMILY participants completing six-month follow-up questionnaire: Intervention II and Control Group 35

V.A.1. Attendance by intervention and participant type 38

V.1.a. Covariates included in impact analyses 42

V.2a. Post-intervention estimated effects using data from six-month follow-up questionnaire to address the primary research questions, Individual Participants....43

V.2b. Post-intervention estimated effects using data from six-month follow-up questionnaire to address the primary research questions, Individual Participants....44

V.2c. Post-intervention estimated effects using data from six-month follow-up questionnaire to address the primary research questions, Couple Participants 44

V.2d. Post-intervention estimated effects using data from six-month follow-up questionnaire to address the primary research questions, Couple Participants 44

V.2e. Post-intervention estimated effects using data from six-month follow-up questionnaire to address the primary research questions, Family Participants 45

V.2f. Post-intervention estimated effects using data from six-month follow-up questionnaire to address the primary research questions, Family Participants 45

V.3. Difference in means between intervention and comparison groups by completion of all outcome response questions methods 46

V.4a. Comparison of outcome 1 (marriage and relationship skills) post-intervention effect size (Cohen's d) among participant types using data from six-month follow-up to address the secondary research questions....48

V.4b. Comparison of outcome 2 (conflict management skills) post-intervention effect size (Cohen's d) among participant types using data from six-month follow-up to address the secondary research questions 48

V.4c. Comparison of outcome 3 (economic stability) post-intervention effect size (Cohen's d) among participant types using data from six-month follow-up to address the secondary research questions 48

B.1. Data used to address implementation research questions – Intervention I 58

B.2. Data used to address implementation research questions – Intervention II 59

B.3. Data used to address implementation research questions – Counterfactual....60

B.4. Key features of the impact analysis data collection....61

C.I. Randomization Assignment by site and participant type 62

Figures

A.1. Logical Model 57

D.1. CONSORT diagram for individual Participants, for studies in which consent occurred before assignment 65

D.2. CONSORT diagram for couple clusters, for studies in which consent occurred before assignment 66

D.3. CONSORT diagram for individual Participants within couples for studies in which consent occurred before assignment 67

D.4. CONSORT diagram for family clusters, for studies in which consent occurred before assignment 68

D.5. CONSORT diagram for individual Participants within families, for studies in which consent occurred before assignment 69

Impact Evaluation of U.S. Committee for Refugees and Immigrants' (USCRI) Refugee Family Strengthening (RFS) Program in 11 Different Program Sites in the U.S.

I. INTRODUCTION

A. Introduction and study overview

This study is the first evaluation of a program developed by the U.S. Committee for Refugees and Immigrants (USCRI) titled Refugee Family Strengthening Program (RFS). The purpose of the RFS program is to strengthen low income refugee and immigrant marriages and relationships with education, mentoring, and comprehensive case management to support long-term relationship success and economic self-sufficiency. The program is designed to help refugees and immigrants enhance their communication and problem-solving skills, and develop financial independence, to improve their marriages, parenting, and family relationships. While the core concepts and skills for Relationship Enhancement (RE®) programs are widely recognized, there are no published evaluations that have documented the effects of the relationship enhancement program for refugees and immigrants.

Over the last three decades more than three million refugees have been admitted to the U.S. through the legal U.S. refugee resettlement process (Krogstad, 2017). These refugees include those from more than 70 different countries (Carey, 2016). They come to the U.S. to escape persecution in their home country for reasons such as race, religion, nationality, or because of their political opinion. In 2017, the United Nations High Commissioner for Refugees (UNHCR) was provided with 75,188 resettlement places, a drop of 54% from 2016, when there were 163,206 places. In 2017 there were 26,782 submissions to the United States, a drop from 108,197 in 2016. In spite of this drop, 38% of all departures (i.e. refugees traveling to the country that has agreed to accept them for resettlement) were to the United States (UNHCR, 2018).

Historically, resettlement work has included public/private partnerships supported by the U.S. government. These partnerships include private and faith-based organizations providing support and opportunities for refugees (Urban Institute, n.d.). The primary purpose for these resettlement programs is to support refugees and help them achieve economic self-sufficiency as quickly as possible after arrival in the U.S. Notably, "economic self-sufficiency" is the only criterion used by the U.S. Department of State to determine resettlement program success. In contrast, other agencies may track other markers of success, but the evaluation of these additional criteria are established by the individual resettlement agencies (Office for Refugee Resettlement, 2015).

The United States State Department (DOS) issues guidelines with basic topics to be covered and service providers design specific programs. A commonality across programs is a focus on how to function in the U.S. where there is an explicit focus on self-sufficiency and the management of expectations. Upon arrival in the U.S., refugees also participate in cultural orientation during their 30 to 90 days of resettlement. The post arrival orientation provides information about a wide range of topics such as enrolling children in school, how to find a job, U.S. laws, health, and public transportation. Thus, the U.S. refugee programs differ, from one organization to the next. However, they share key commonalities, such as an emphasis on the importance of employment and the goal of helping refugees make a successful transition to their new home country.

The RFS program was USCRI's effort to provide additional programming that maximized the chances that refugee families will pull together and operate as a cohesive unit, adapt to living in a new and different country, and be successful. The RFS program targets specific goals that provide (1) families with pathways to financial self-sufficiency, (2) emotional support for marriage/family relationships, (3) conflict management skills, and (4) healthy bonds between parents and children. These goals underscore the importance of becoming contributing members of society here in the U.S.

The RFS program is based on the RE® program, developed by Guerney (1977). Accordino and Guerney (2002) point to strong evidence of the effectiveness of RE® in improving general adjustment level and quality of relationships. Additionally, the RFS program was informed by related work developed specifically for refugees and immigrants (Guerney, Ortwein, & Amin, 2007; Guerney, Ortwein, & Amin, 2011) a financial literacy component (Guerney, Ortwein, Amin, Chaterdon, & Solly, 2008) and additional adaptations made by Ortwein (2014), specifically working with USCRI.

Review of the Literature

Dion (2005) highlighted healthy marriage programs and indicated a number of studies had examined the effects of such programs and found that they improved relationship satisfaction and communication among romantically involved couples. Her concern was that in the programs tested, the participating couples were largely white, well-educated, middle class, and engaged or already married. Thus, there are questions as to whether such programs will be effective with disadvantaged populations. Dion's work was published in 2005, but the lack of research involving non-white, disadvantaged populations, including refugee and immigrant groups, continues to be a concern.

In a 2008 meta-analysis of 117 published research studies, Hawkins and colleagues (2008) examined the effects of marriage and relationship education. These researchers found that such programs did produce significant improvement in relationship quality and communication skills. There were no gender differences relative to program impact. The researchers did, however, note a lack of racial/ethnic and economic diversity in the studies they examined. This lack of

racial/ethnic and economic diversity prevents us from generalizing the results of this meta-analysis to other groups, like refugee and immigrant populations in the U.S.

In 2012, Johnson (2012) addressed the federal initiatives designed to promote healthy marriage, initiatives which have specifically targeted poor couples and couples of colors. While such programs have shown some degree of success among white and middle-class couples, Johnson noted that the results, so far, in applying these interventions to the couples targeted by these federal initiatives, have been disappointing. In contrast, Amato (2014) suggested, "contrary to the notion that disadvantaged couples do not benefit from relationship education, these couples may be the main beneficiaries of these services, provided they are able to keep their unions intact," (p. 353). Amato's comments were from his report based on data from the 15 month, Building Strong Families project.

Unlike, the RFS program evaluated, which is classroom based, and involved large numbers of participants across program sites, Bellinger (2013) reported results from a clinician working directly with one couple. Bellinger described how a program that applied the Relationship Enhancement model was of benefit to a Bhutanese couple. Like many refugees who are resettled in Western countries, the couple experienced marital tension and threats to traditional gender roles. The program helped the couple identify their core concerns about the changes in their lives and helped them address new responsibilities related to household chores, parenting, employment, and finances.

Jackson and colleagues (2016) asked Black, Latino and White newlywed spouses living in low-income neighborhoods about stressors in their marital relationships. The researchers first asked participants to list their concerns and then were asked to complete a standard relationship problem inventory, rating each item as to its degree of severity in their relationship. Relationship problems (e.g. communication) were rated as severe on the inventory; however, factors external to the relationship (e.g. children) were more likely to be identified on the list. The researchers concluded that efforts to strengthen marriages among low-income couples may well be more effective if, in addition, to efforts to address relationship problems, external stressors are addressed. This might take the form of childcare, financial assistance, or job training.

In another 2016 study, Williamson and others examined data from 1,034 low-income couples participating in a Supporting Healthy Marriage project. The researchers found that at 30 months after randomization, intervention couples reported a higher relationship satisfaction than did control couples, regardless of their level of risk at pretreatment. Among higher risk couples the intervention also improved communication. However, treatment effects on satisfaction were not mediated by level of communication and improvements in communication did not translate into improved satisfaction.

In addition to Amato (2014) who indicated that these programs have the potential to be of real benefit to low-income/disadvantaged participants, Hawkins (2019) also indicated these programs hold promise. In his review of federally funded healthy marriage programming, Hawkins noted scholars and other professionals have taken a negative view of such programs and have been ready to write the effort off as a failure. After acknowledging that others disagree with him,

Hawkins examined the research. He indicated "Large scale, rigorous, randomized control trial evaluation studies reveal promising successes, disappointing failures, and nuanced findings." He noted that participants in healthy marriage programs reported enjoying the program and indicated programs were of help. His review of evaluations also found rigorous impact research has shown mixed results. Some studies have shown small effects on helping distressed, low-income married couples increase their commitment and remain married. There is no evidence to show that couples completing the program are more likely to get married, however, such programming may help some couples stay together longer. Additionally, some studies have shown couples can learn to reduce destructive conflict and physical and emotional abuse. There is also evidence that programming helps couples improve communication, understanding, and co-parenting. Hawkins also noted that the most disadvantaged and distressed couples are the ones who benefit the most, and children of participating couples have also been shown to benefit.

Compared to the amount of research concerning the effects of relationship education/healthy marriage programs on participants who are white and middle class, research demonstrating the effectiveness of relationship education/healthy marriage programs among low-income, disadvantaged, minority populations are still relatively limited. Research concerning the effects of such programs among refugee populations is even harder to find. Recently, Shaw and Funk (2019) examined research on social service programs implemented with refugees worldwide. They began with 1,402 resources, but only 68 of these met their review criteria and were included in their analysis. Of the 68 studies, 57 were conducted in high income countries. Some of the programs were concerned with general adaptation (n=13), others with relationships (n=20), or financial or employment support (n=15). The authors indicated only six of the 68 studies used pretest-posttests and only one used an "experimental design." The authors indicated they used the Critical Appraisal Skills Program (CASP) checklist (CASP, 2018) to examine the level of rigor across study methods. Of studies that included quantitative methods, one randomized controlled trial (Stark et al., 2018) and one case control study (Mansson & Delander, 2017) were the only studies that included a control condition. The randomized control trial was a large sample study concerned with empowerment of refugee girls in Ethiopia. It was one of the few studies the authors reviewed that reported a negative outcome. The other study dealt with a mentoring program in Sweden.

Well-controlled studies, i.e. studies that have used a control group, and especially randomized control trials (RCTs), that have examined relationship enhancement type programs in refugee populations are almost non-existent. The present study makes an important contribution to the literature by using a randomized control trial design to examine the effects of a relationship enhancement program provided to refugees who have resettled in the United States.

USCRI applied expertise and culturally and linguistically appropriate resources developed through the RFS program at selected field sites and with partner agencies in the following 11 cities across the U.S: Albany, NY, Colchester, VT, Dearborn, MI, Des Moines, IA, Erie, PA, Houston, TX, Kansas City, MO, Los Angeles, CA, Philadelphia, PA, Raleigh, NC, and St. Louis, MO. Although there is a history of established agencies providing resources and programs for refugees and immigrants, there is a lack of evaluation studies indicating the effectiveness of programs tailored to healthy marriage and relationships.

The purpose of the study was to determine if there were long-term positive effects of a healthy relationship program for refugees and immigrants. This report provides a detailed guide of how the program was implemented and data was collected, the results from the detailed analysis, information learned from implementation, and what key findings suggest for implementing similar programs in the future. The program designed outcomes in the logic model addressing the priority areas: Marriage and relationship skills, conflict management skills, and economic stability. The study examined whether two interventions based on relationship enhancement, conflict resolution, and economic stability produced significant improvements for participants in three outcome measures, when compared to controls, at the six-month follow-up. The study used a randomized control trial (RCT) design.

B. Primary research question(s)

Research Questions for participants who received Intervention I: Relationship Enhancement® and Conflict Resolution program components by participant group

Research Questions for Individual Participants:

1. Did *individual* participants demonstrate a greater improvement in marriage and relationship skills compared to the control group at the six-month follow-up?
2. Did *individual* participants demonstrate a greater improvement in conflict management skills compared to the control group at the six-month follow-up?
3. Did *individual* participants demonstrate a greater improvement in economic stability compared to the control group at the six-month follow-up?

Research Questions for Couple Participants:

1. Did *couple* participants demonstrate a greater improvement in marriage and relationship skills compared to the control group at the six-month follow-up?
2. Did *couple* participants demonstrate a greater improvement in conflict management skills compared to the control group at the six-month follow-up?
3. Did *couple* participants demonstrate a greater improvement in economic stability compared to the control group at the six-month follow-up?

Research Questions for Family Participants:

1. Did *family* participants demonstrate a greater improvement in marriage and relationship skills compared to the control group at the six-month follow-up?
2. Did *family* participants demonstrate a greater improvement in conflict management skills compared to the control group at the six-month follow-up?
3. Did *family* participants demonstrate a greater improvement in economic stability compared to the control group at the six-month follow-up?

Research Questions for participants who received Intervention II: Relationship Enhancement® and Financial Management program components by participant group

Research Questions for Individual Participants:

1. 10) Did *individual* participants demonstrate a greater improvement in marriage and relationship skills compared to the control group at the six-month follow-up?
2. 11) Did *individual* participants demonstrate a greater improvement in conflict management skills compared to the control group at the six-month follow-up?
3. 12) Did *individual* participants demonstrate a greater improvement in economic stability compared to the control group at the six-month follow-up?

Research Questions for Couple Participants:

1. 13) Did *couple* participants demonstrate a greater improvement in marriage and relationship skills compared to the control group at the six-month follow-up?
2. 14) Did *couple* participants demonstrate a greater improvement in conflict management skills compared to the control group at the six-month follow-up?
3. 15) Did *couple* participants demonstrate a greater improvement in economic stability compared to the control group at the six-month follow-up?

Research Questions for Family Participants:

1. 16) Did *family* participants demonstrate a greater improvement in marriage and relationship skills compared to the control group at the six-month follow-up?
2. 17) Did *family* participants demonstrate a greater improvement in conflict management skills compared to the control group at the six-month follow-up?
3. 18) Did f*amily* participants demonstrate a greater improvement in economic stability compared to the control group at the six-month follow-up?

This study is registered with ClinicalTrials.gov and the Identifier is NCT02829086.

C. Secondary research question(s)

1. There is one secondary research question: Do the effects of the intervention (Intervention I vs. control or Intervention II vs. control) differ by type of participant (individual, couple, or family)?

II. INTERVENTION AND COUNTERFACTUAL CONDITIONS

This section provides a brief description of the intervention being evaluated and the services that are intended as the comparison to the intervention, including any "social services/resources/support as usual" resources available. Participants assigned to the two intervention groups received curriculum sessions and case management services. Participants assigned to the control group received case management services. Case management services were provided on an as-needed basis and the hours received by participant may vary.

A. Description of program as intended

a. **Intended components:** The RFS Program provided participants assigned to Interventions I and II with relationship education, comprehensive case management, and support services with the idea/goal of supporting long-term relationship success and economic self-sufficiency. Program delivery did not differ by participant type (*individual*, *couple*, or *family* participant) and sessions were delivered in groups with mixed participant types. Family members were served together.

b. **Intended content:** The 16-hour intervention utilized the curriculum Relationship Enhancement® and two supplementary curricula: Family Stress and Conflict Management for Intervention I and Financial Management for Intervention II. The curricula addressed the four major topic areas guiding the research study: Relationship Skills, Parent-Child Relationship, Conflict Management, and Financial Management. Participants received eight hours of Relationship Enhancement® (Intervention I and II), eight hours of Family Stress and Conflict management (Intervention I), and eight hours of Financial Management (Intervention II). Table II.1 provides the details of the topics covered for each curriculum.

c. **Planned dosage and implementation schedule:** The 16-hour curriculum-based intervention was delivered in either, (1) four-hour sessions once per week, (2) two-hour sessions once a week, and (3) one-hour sessions once a week to allow flexibility for scheduling. Participants could not meet more than four hours per week total. Sessions occurred on these schedules until all 16 hours of the curriculum are implemented. The dosage was delivered across 11 program sites across the United States. Curricula were implemented sequentially, and lessons were not skipped or removed.

d. **Intended delivery:** Participants received the intervention in small group sessions at designated times and locations within their communities. Sites were selected for ease of travel and access for participants. Examples of sites for group sessions included RFS program offices, local community centers, and participants' homes. If a participant was unable to attend or missed a session, they had the opportunity to "make up" through a one-on-one session with the facilitator. In addition to receiving the curricula, participants in the intervention group received case management services. This included, but was not limited to, developing a support service plan, linkages for social adjustment services, employment assistance, health and medical referrals, and educational support. Participants were enrolled in the group on a rolling basis and would start programming when a cohort was formed. Groups were not separated by participant type, for example, individual participants would attend group classes with couples and family participants. Couples and families were enrolled in the same cohort and would attend group sessions together. The recommended group size was between 6 and 12 individuals. Program coordinators generally followed this but occasionally there were fewer or more participants. There also was no set mixture of individuals, couples, and families. When applicable, program coordinators and

program facilitators served participants and facilitated curriculum activities in the native language of the participants. For example, one of the program coordinators taught in Spanish when serving Spanish-speaking participants. Other program coordinators taught in Arabic when serving Arabic-speaking participants. When this was not possible, interpreters were utilized for all activities on an as-needed basis. In the case that several languages were spoken in a group, there would be more than one interpreter present during the group classes providing interpretation for the facilitator. All print materials were available to participants in their native language. The 15 most frequently spoken languages by the participants were: Amharic, Burmese, Chin, Dari, Farsi, French, Karen, Kinyarwanda, Korean, Nepali, Pashto, Somali, Spanish, Swahili, and Tigrinya.

e. **Target population:** The population targeted was self-identified refugee and immigrant couples, families, and individuals, ages 16 and older.

f. **Education and training of staff:** USCRI recruited Program Coordinators and Case Managers with experience in providing services to refugee and immigrant communities and many spoke the languages of their clients. Table II.2. provides details specific to the educational and training requirements of staff that deliver curriculum components and resources. A Bachelor's Degree was an educational requirement for both Program Coordinators and Case Managers. RFS program staff included men and women of different ages and backgrounds who were familiar with the needs of participants within the community. All facilitators that delivered curriculum components received 24 hours of training in curriculum delivery from the Institute for the Development of Emotional and Life Skills (IDEALS) curriculum developer. In addition to curriculum-training, all RFS Program staff completed training in ethics and human subjects research provided through Protecting Human Research Participants (PHRP) endorsed by National Institute of Health (NIH). The RFS program staff completed ethics and human subjects research re-certification every three years and received booster and "refresher" courses trainings on the curriculum components each year.

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

Component	Curriculum and content	Dosage and schedule	Delivery	Target Population
Intervention I				
Relationship Skills workshops	Introduction to Relationship Enhancement curriculum: Understanding partner's perspectives; avoiding destructive conflict; and communicating effectively.	Eight hours delivered in either, (1) four-hour sessions once per week, (2) two-hour sessions once a week, and (3) one-hour sessions once a week to allow flexibility for scheduling. Participants could not meet more than four hours per week total.	Group lessons provided at the program office or at a designated location in the community by one trained facilitator, either the Program Coordinator or Case Manager at each of the 11 sites. Interpreters were utilized on an as-needed basis.	Participants were self-identified as immigrants or refugees aged 16 or older. They participated as individuals, couples, or families.
Family Stress and Conflict Management workshops	Family Stress and Conflict Management curriculum: Communication skills, emotional regulation, stress management skills.	Eight hours delivered in either, (1) four-hour sessions once per week, (2) two-hour sessions once a week, and (3) one-hour sessions once a week to allow flexibility for scheduling. Participants could not meet more than four hours per week total.	Group lessons provided at the program office or at a designated location in the community by one trained facilitator, either the Program Coordinator or Case Manager at each of the 11 sites. Interpreters were utilized on an as-needed basis.	Participants were self-identified as immigrants or refugees aged 16 or older. They participated as individuals, couples, or families.
Case Management	Support service plan, links to social adjustment services, employment assistance, health and medical referrals, and educational support.	No assigned number of case management hours for participants. Services are delivered on an as-needed basis.	One-on-one appointments with Case Managers or Program Coordinators. Interpreters were utilized on an as-needed basis.	Participants were self-identified as immigrants or refugees aged 16 or older. They participated as individuals, couples, or families.
Intervention II				
Relationship Skills workshops	Same as Intervention I	Same as Intervention I	Same as Intervention	Same as Intervention I
Financial management workshops	Financial Management curriculum: budgeting skills, money management.	Same as Intervention I	Same as Intervention	Same as Intervention I
Case Management	Same as Intervention I	Same as Intervention I	Same as Intervention	Same as Intervention I
Counterfactual				
Case Management	Same as Intervention I and II	Same as Intervention I and II	Same as Intervention I and II	Same as Intervention I and II

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

Component	Education and initial training of staff	Ongoing training of staff
Interventions I and II		
Relationship Skills workshops	Prior experience with social services provision, employment services or refugee resettlement is preferred, and all staff must have dedication to the human rights of refugees, immigrants, asylum seekers and displaced people. All facilitators receive 24 hours of face-to-face training over the course of three days from the Curriculum Developer IDEALS and Program Officers prior to starting workshops.	Each year USCRI held an annual meeting which includes a dedicated curriculum session for opportunities to "refresh" and to discuss any question or concern about the curriculum or curriculum implementation.
Family Stress and Conflict Management workshops	Prior experience with social services provision, employment services or refugee resettlement is preferred, and all staff must have dedication to the human rights of refugees, immigrants, asylum seekers and displaced people. All facilitators receive 24 hours of face-to-face training over the course of three days from the Curriculum Developer IDEALS and Program Officers prior to starting workshops.	Each year USCRI held an annual meeting which includes a dedicated curriculum session for opportunities to "refresh" and to discuss any question or concern about the curriculum implementation.
Financial Management workshops	Prior experience with social services provision, employment services or refugee resettlement is preferred, and all staff must have dedication to the human rights of refugees, immigrants, asylum seekers and displaced people. All facilitators receive 24 hours of face-to-face training over the course of three days from the Curriculum Developer IDEALS and Program Officers prior to starting workshops.	Each year USCRI held an annual meeting which includes a dedicated curriculum session for opportunities to "refresh" and to discuss any question or concern about the curriculum implementation.
Counterfactual		
Case Management	Prior experience with social services provision, employment services or refugee resettlement is preferred, and all staff must have dedication to the human rights of refugees, immigrants, asylum seekers and displaced people.	Monthly meetings with Program Officers and Program Coordinators, site visits, USCRI orientation and trainings.

B. Description of counterfactual condition as intended

This counterfactual condition is really "social services/resources/support as usual" Control group participants had access to case management services only. After control participants participated in all data collection activities associated with being a member of the control group – pretest, posttest, and six-month follow-up - they were given an opportunity to participate in RE® classes. Participants assigned to the control group received case management services. Case management services were provided on an as-needed basis and the hours received by participant may have varied.

1. **Intended components:** The RFS Program provided participants assigned to the Control Group with comprehensive case management only.
2. **Intended content:** This included, but was not limited to, developing a support service plan, linkages for social adjustment services, employment assistance, health and medical referrals, and educational support. Control participants were monitored through a minimum of monthly periodic contact via phone and scheduled appointments to offer these services.

3. **Planned dosage and implementation schedule:** Control group participants received case management services throughout their participation in the study (approximately six months). Same as the intervention group participants, there was not an assigned number of case management hours for participants. Due to case management being based on need, some did not receive any hours and others received numerous hours of support.
4. **Intended delivery:** Control participants met with Program Coordinators and Case Managers at RFS Program sites to complete data collection and meet with their assigned case managers. Control participants scheduled appointments for services as needed. Similar to the intervention groups, interpreters were utilized on an as-needed basis.
5. **Target Population:** The population targeted was the same as the intervention groups; self-identified refugee and immigrant couples, families, and individuals, ages 16 and older.
6. **Education and training of staff:** USCRI utilized RFS Program staff to engage control participants in case management appointments and data collection. The required education and training of staff was the same for the intervention and control groups. RFS Program staff completed training in ethics and human subjects research provided through PHRP endorsed by NIH. The RFS program staff completed ethics and human subjects research re-certification every three years.

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

The evaluation examined data collected across five implementation elements identified in Table II.3 to determine which, if any, of the elements impacted how the program was received by the participants. The aim of the examining the implementation elements was to learn if the program was delivered consistently across each of the intervention groups, how the participants perceived the quality of the program delivery, if dosage differed between the intervention groups, and if there were any contextual factors that affected implementation. These questions were selected with the intent to integrate findings together with those of the impact evaluation. The research questions examined if there were any differences in implementation between each Intervention Group and Control Group and to determine if differences in implementation occurred that may have affected the impact of the program across the groups.

Table II.3. Research questions of the implementation study, by implementation element

Implementation element	Research question
Intervention Group I Questions	
Fidelity	Were all intended intervention components of Relationship Enhancement® and Family Stress and Conflict Management delivered as proposed during curriculum training? Who facilitated each of the RFS workshop sessions across the 11 sites?
Dosage	What was the average amount of curriculum hours that participants received?
Quality	What was the quality of RFS program staff-participant interactions?
Engagement	How engaged were intervention group members in the program?

Implementation element	Research question
Context	What, if any, additional programming and resources were made available to participants? What external factors affected implementation?
Intervention Group II Questions	
Fidelity	Were all intended intervention components of Relationship Enhancement® and Financial Management delivered as proposed during curriculum training? Who facilitated each of the RFS workshop sessions across the 11 sites?
Dosage	What was the average amount of curriculum hours that participants receive?
Quality	What was the quality of RFS program staff-participant interactions?
Engagement	How engaged were intervention group members in the program?
Context	What, if any, additional programming and resources were made available to participants? What external factors affected implementation?
Control Group Questions	
Fidelity	What content did the control group receive from any sources during the evaluation period? Who delivered any services similar to the intervention to the control group members during the evaluation period?
Dosage	How often did the control group members participate in services similar to the intervention from other sources, on average?
Quality	What was the quality of staff–participant interactions in services similar to the intervention received from other sources, if known?
Engagement	How engaged were control group members in services similar to the intervention received from other sources, if known?
Context	What, if any, programming and resources were made available to participants?

III. STUDY DESIGN

This section provides a description of the study design and the process for creating intervention and comparison groups. The study is an RCT consisting of three groups: Intervention I, Intervention II, and Control, as described in Section B. Data collection for all three groups occurred at baseline (pretest) – right before the two intervention groups began participating in the Relationship Enhancement classes; posttest – immediately after the two intervention groups finished the Relationship Enhancement classes; and follow-up – three months after completion of the posttest (approximately six months from the pretest).

A. Sample formation and research design

Recruitment. Eligible participants were self-identified as immigrants or refugees aged 16 or older. They participated as individuals, couples, or families. Program Coordinators did most (if not all) of the recruitment. Sometimes this took place because potential participants came into their agencies for related services and were recruited to the program. Other times program coordinators went out to the community and distributed materials. Often referrals came from other service providers in the same agency. Recruitment included distributing fliers and brochures; other times it involved speaking to participants in other services such as orientation classes, workshops, and programs.

Randomization. Study participants fell into three groups: *individuals*, *couples*, and *family*. Within each group, the same randomization strategy was used. To be clear, in no case was one member of a *couple* or *family* assigned to Intervention I and another member assigned to Intervention II or control. Both members of a *couple*, and all members of a *family* were assigned to the same group.

Randomization was conducted on a rolling basis. Prior to randomization, participants provided written consent to participate in the study and completed baseline data collection. Participants were aware they would be randomly assigned to a group that includes classes or a group that does not include classes but did include case management. After completing the consent forms, participants completed a questionnaire to collect demographic information named the Applicant Characteristic Questionnaire. Participants were then randomly assigned to one of the three treatment conditions using an envelope system designed to randomly assign participants to groups. To assure that participants were randomly assigned into one of the three respective groups without bias, sealed security envelopes were sequentially numbered (001, 002, 003, 004, etc.) within each of the three participating groups. There were three separate sequences, one for each group: *individual, couples,* and *family*. Participants opened the sealed security envelope. In each security envelope was a 3x5 index card with the group code (Intervention I, Intervention II or Control) written on the card, thus revealing the group to which the participant(s) had been randomly assigned. This way, the Program Coordinators or Case Managers did not have any influence as to the assignment of participants, and both the Program Coordinator/Case Manager and the participant(s) learned of the group assignment at the same time. Following the opening of the envelope, the assigned group was recorded on the master list copy at the program site. The program staff, including Case Managers, Data Collection Specialists and Program Coordinators were trained and informed that group selection cannot be changed. Program staff informed participants that group selection cannot be changed.

The envelopes were created, sorted, and sealed by the Principal Investigator. At the end of each quarter, the number of new enrollees in each group was checked for consistency in assignment and to see if the sites are maintained a 1:1:1 ratio. The intended probability was for each *individual* participant, *couple*, or *family* to have a one in three probability of assignment to a specific group. Randomization was stratified by site and participant type (see Appendix C). The sizes of each group were in roughly the expected proportion. Participants completed the baseline questionnaire, or pretest, prior to group assignment. After the completion of the questionnaire, those who were assigned to an intervention were given the specifics as to dates and times of classes they would be attending. Participants who were assigned to the control group were given the date/time for their case management appointment. Staff members made sure that the time/date scheduled worked for the participant. If the participant(s) indicated they could not make the assigned class/appointment the staff member worked with the participant to secure a workable date/time for the class/appointment.

Consent Process. IRB approval was obtained through The Western Institutional Review Board (WIRB). Persons age 16 years and older were allowed to participate in this study. The study required written participant consent. When the prospective participant was under 18, then active parental consent was required (i.e. parent or guardian must sign giving permission for their child to participate), as was participant consent. Consent was obtained before participants took part in any aspect of the program and before they completed the baseline questionnaire. Prospective participants were informed of the study the first time they met with the Program Coordinator or Case Manager at intake. Prior to beginning the intake process, these prospective participants were provided with a detailed informed consent form which explained the aims of the program and any benefits they may have expected from participation in the study as well as any potential risks related to participation.

Risk associated with participation in the project was minimal. Participants may have felt uncomfortable responding to some items in the questionnaire because some topics may be viewed as personal or involve sensitive issues. Participants were informed that responses to questionnaire items were voluntary. Prospective participants who decided to not participate in the study still received case management services.

B. Data collection

This section provides a description of the additional data sources and data collection methods used to address the implementation study research questions. Data collection was ongoing throughout the program. To examine implementation elements, the evaluation looked at the information management system developed for grantees and data gathered by RFS Program Officers for ongoing site monitoring. The evaluation also reviewed qualitative data gathered through focus groups that occurred at selected sites between 2016 and 2020. Tables B.1, B.2, and B.3 in Appendix B provide detailed information for each data source examined for each implementation element by group.

The information management system developed for use by OFA grantees, named nFORM, was used to track all programmatic activities, including but not limited to enrollment, attendance, program completion, and session delivery. nFORM was also used to note service contacts with

participants. Participants' information was logged in nFORM including demographic and contact information. Three questionnaires were housed in the nFORM system that provided data to the evaluation. These included The Applicant Characteristic Questionnaire, nFORM Exit Survey, and nFORM Entrance Survey. Data were exported by the evaluation on a quarterly basis from nFORM.

Focus groups conducted included questions that explored the following themes: What are the strengths of the RFS program? What are the challenges of the RFS program? What recommendations do participants have to strengthen the program? Are there other issues participants believe should be addressed by RFS program staff for future programming efforts? Additionally, when needed, interpreters were present during the groups to ensure that participants had every opportunity to share their experiences of the program. A member of the local evaluation team was assigned to conduct all of the group interviews and to summarize the findings.

Between July 2016 and September 2019, the USCRI Program Officers selected two to three sites to conduct site monitoring visits. During these visits, the Officers observed workshop sessions to examine the facilitator's skills and participant engagement. The Officers also conducted individual client interviews and reviewed case management materials such as case notes and attendance sheets. The Program Officers completed a site monitoring report for each site after their visit. The report included the Program Officer's observations and recommendations for improvement.

The USCRI Program Officers monitored the implementation activities through operational reports in nFORM and with bi-monthly team meetings and monthly check-ins with each site. During these meetings, Program Coordinators provided programmatic updates and reported any external events that affected implementation. The evaluators anticipated learning if there were any differences in contextual factors across sites that may have affected implementation. The evaluators also hoped to learn if the programmatic components were delivered consistently across program sites.

Participants in the two intervention groups completed the pretest questionnaire prior to participating in any classes. Participants took the posttest questionnaire after all sessions were completed. Generally, this was on the last day of sessions, at the completion of that class. Participants in the control group completed their pretest and posttest questionnaires in the same time frame as the two intervention groups as a simple way to keep track of the timing of questionnaire completion. Generally, the pretest was completed at their intake/enrollment appointment. For control participants, Program Coordinators, Case Managers and Data Collection Specialists at each program site scheduled an appointment with the control participants to complete the posttest within a two-week time period from the last day of sessions. All participants (from both intervention groups and the control group) completed the follow-up questionnaire three months after completion of the program, about six months after completion of the pretest. For the intervention groups, a date was designated for participants to complete the follow-up questionnaire. Table B.4 in Appendix B provides the details of the timing of data

collection across all groups. For control participants, case management appointments were made to coincide with the six-month follow-up time. Participants who did not show up for a scheduled data collection date were contacted by phone, email, through friends, or in-person to schedule another date/time for them to come complete their questionnaire. When participants indicated a willingness to complete the questionnaire but did not come to the program site to complete the questionnaire, project staff took the questionnaire to them. Completion of the questionnaire was voluntary, but those who did complete it received a gift card of a nominal amount for each of the three local evaluation questionnaires they were asked to complete.

All project staff were trained in evaluation procedures. This included Data Collection Specialists (DCS) who participated in training with specific attention given to data collection protocols. Whenever possible, participants completed their questionnaires electronically, using Samsung Galaxy tablets made available for use at the USCRI project sites. Participants completed all evaluation questionnaires via Qualtrics web-based software. The DCS or RFS program staff member inputted the following for each participants' questionnaire: site name, the participant/couple/family member's nFORM ID number, test time (pretest, posttest, follow-up), group code, and test date. The participant code number was the same number for all three questionnaires that a given participant completed. This then allowed the evaluation team to match participants' initial pretest questionnaire responses, with the same participant's responses to the posttest and follow-up.

The program staff made several efforts to ensure that participants were able to complete the questionnaires with ease. For each data collection point, a paper copy of the questionnaire was made available to each participant in their native language, as needed. In addition to English, the questionnaire was available in the 15 most frequently spoken languages by the those served by USCRI. WIRB translated the consent forms and questionnaires. These documents were reviewed by RFS program staff and interpreters were asked to identify any issues or errors, and to provide feedback for ease of use. All feedback and revisions were directed to WIRB, which issued final translations of the questionnaires. The printed questionnaire was designed to look exactly like the Qualtrics questionnaire, so participants were able to view the paper questionnaire alongside the corresponding questionnaire in English on the tablet and enter their responses electronically. They also had the option of completing the hardcopy (paper) version of the questionnaire. The DCS aided as needed. Paper questionnaires were offered to participants if they did not feel comfortable utilizing the tablets, did not have access to the internet, or desired to complete the questionnaire in their native language. When participants completed the paper questionnaire, the questionnaires were collected and secured by the DCS or RFS Program staff member who then entered the data from the paper questionnaire into Qualtrics. For participants that presented with literacy challenges, interpreters were present to provide participants assistance in understanding questionnaire instructions, as needed. Sometimes this involved the interpreter reading the questions to the participant.

To minimize problems in completing and submitting questionnaires, the DCS or RFS Program staff member used a tablet to demonstrate the following: how to answer questions, how to move to the next page, and how to submit the completed questionnaire. DCSs maintained regular

contact with all participants at their site with phone calls or other contact at least once per month to “check in” and gather new/additional contact information as needed. Participants were encouraged to remain in contact to receive incentives for completing the posttest and follow-up questionnaires. Because all participants, from both intervention groups and the control group, were receiving comprehensive case management services, other project staff were in periodic contact with participants and also encouraged them to complete posttest and follow-up questionnaires.

IV. ANALYSIS METHODS

This section provides a description of the data analyses focusing on three composite outcome measures by participant type (*individual*, *couple*, and *family*). Primary and secondary research questions were evaluated to determine impacts on the outcome measures. The outcome measures, data preparation, sample, and approach are detailed below.

A. Analytic sample

Separate analyses were performed to compare six-month follow-up effects to the baseline for each participant type (e.g. *individual*, *couple*, and *family*) therefore three separate analytic samples were established. To compare program effects, only participants who completed both the baseline assessment and the six-month follow-up questionnaire were included. For the analytic sample of *individual* participants, a maximum of 939 treatment group participants are included (Intervention I: n= 469, Intervention II: 470), corresponding to the number of treatment group participants who completed both the baseline and follow-up questionnaire. The number of *individual* treatment group participants in the analytic sample differed by outcome (outcome 1: Total = 938, Intervention I = 468, Intervention II = 470; outcome 2: Total = 939, Intervention I = 469, Intervention II = 470; outcome 3: Total = 937, Intervention I = 468, Intervention II = 469) See Tables IV.1a and IV.1b below. The *individual* participant analytic sample also included maximum of 466 control group participants, which corresponds to the number of control group participants that completed the baseline questionnaire and follow-up questionnaire. The number of control group participants in the analytic sample of *individual* participants varied by outcome (outcome 1 = 466; outcome 2 = 465; outcome 3 = 465) as detailed in Tables IV.1a and IV.1b.

Table IV.1a. Individual sample sizes by intervention status: Intervention I and Control Group Participants

Number of individuals	Intervention I sample size	Comparison sample size	Total sample size	Total response rate	Intervention response rate	Comparison response rate
Assigned to condition	649	649	1298	N/A	N/A	N/A
Contributed a baseline survey	509	509	1018	78.4%	78.4%	78.4%
Contributed to follow-up survey (timing)	469	466	935	72.0%	72.3%	71.8%
Contributed to follow-up (timing) (accounts for item non-response and any other analysis restrictions)						
Outcome 1	468	466	934	72.0%	72.1%	71.8%
Outcome 2	469	465	934	72.0%	72.3%	71.6%
Outcome 3	468	465	933	71.9%	72.1%	71.6%

Note: Individuals were randomized after consent to participate was acquired.

Table IV.1b. Individual sample sizes by intervention status: Intervention II and Control Group Participants

Number of individuals	Intervention II sample size	Comparison sample size	Total sample size	Total response rate	Intervention response rate	Comparison response rate
Assigned to condition	654	649	1303	N/A	N/A	N/A
Contributed a baseline survey	501	509	1010	77.5%	76.6%	78.4%
Contributed to follow-up survey (timing)	470	466	936	71.8%	71.9%	71.8%
Contributed to follow-up (timing) (accounts for item non-response and any other analysis restrictions)						
Outcome 1	470	466	936	71.8%	71.9%	71.8%
Outcome 2	470	465	935	71.8%	71.9%	71.6%
Outcome 3	469	465	934	71.7%	71.7%	71.6%

Individuals were randomized after consent to participate was acquired.

For the analytic sample of *couple* participants, 339 treatment group clusters (Intervention I: 175, Intervention II: 164), equating to 666 treatment group participants (Intervention I: 343, Intervention II: 323) are included (see Tables IV.1c and IV.1d). This sample includes treatment group participants who completed both the baseline and follow-up questionnaire, couples where only a single member completed the baseline questionnaire were excluded from the sample. The analytic sample of *couple* treatment group participants did not vary by outcome. A total of 168 control group clusters (322 participants) are included in the follow-up analysis of *couple* participants, which corresponds to the number of control group participants that completed the baseline questionnaire and follow-up questionnaire. The analytic sample of *couple* control group participants varied by outcome (outcome 1: control = 322; outcome 2: control = 322; outcome 3: control = 321) as detailed in Tables IV.1c and IV.1d.

Table IV.1c. Cluster and individual sample sizes by intervention status: Couple Participants, Treatment I, and Control Group

Number of:	Intervention sample size	Comparison sample size	Total sample size	Total response rate	Intervention response rate	Comparison response rate
Clusters						
Clusters: At beginning of study	202	197	399	N/A	N/A	N/A
Clusters: Contributed at least one individual at baseline	183	178	361	90.5%	90.6%	90.4%
Clusters: Contributed at least one individual at second follow-up (timing)	175	168	343	86.0%	86.6%	85.3%
Individuals in non-attriting[a]						
Individual: At time that clusters were assigned to condition	361	350	711	N/A	N/A	N/A
Individual: Contributed a baseline survey	361	350	711	100.0%	100.0%	100.0%
Individual: Contributed to second follow-up survey (timing)	343	323	666	93.7%	95.0%	92.3%
Individual: Contributed to the impact analysis at second follow-up (timing) (accounts for item non-response and any other analysis restrictions)[b]						
Outcome 1	342	322	664	93.4%	94.7%	92.0%
Outcome 2	342	322	664	93.4%	94.7%	92.0%
Outcome 3	342	321	663	93.2%	94.7%	91.7%

Note: Individuals were randomized after consent to participate was acquired.

Table IV.1d. Cluster and individual sample sizes by intervention status: Couple Participants, Treatment II and Control Group

Number of:	Intervention sample size	Comparison sample size	Total sample size	Total response rate	Intervention response rate	Comparison response rate
Clusters						
Clusters: At beginning of study	196	197	393	N/A	N/A	N/A
Clusters: Contributed at least one individual at baseline	173	178	351	89.3%	88.3%	90.4%
Clusters: Contributed at least one individual at second follow-up (timing)	164	168	332	84.5%	83.7%	85.3%
Individuals in non-attriting clusters[a]						
Individual: At time that clusters were assigned to condition	345	350	695	N/A	NA	N/A
Individual: Contributed a baseline questionnaire	345	350	695	100.0%	100.0%	100.0%
Individual: Contributed to second follow-up questionnaire (timing)	323	323	646	92.9%	93.6%	92.3%
Individual: Contributed to the impact analysis at second follow-up (timing) (accounts for item non-response and any other analysis restrictions) [b]						
Outcome 1	323	322	645	92.8%	93.6%	92.0%
Outcome 2	323	322	645	92.8%	93.6%	92.0%
Outcome 3	323	321	644	92.7%	93.6%	91.7%

Individuals were randomized after consent to participate was acquired.

For the analytic sample of *family* participants, 365 treatment group clusters (Intervention I: 179, Intervention II: 186), equating to a maximum of 1,193 treatment group participants (Intervention I: 599, Intervention II: 594) are included (see Tables IV.1e and IV.1f). This sample includes treatment group participants who completed both the baseline and follow-up questionnaire, families where one or more members failed to complete the baseline questionnaire were excluded from the sample. The analytic sample of *family* treatment group participants varied by outcome (outcome 1: Total = 1,192, Intervention I = 598, Intervention II = 594; outcome 2: Total = 1,192, Intervention I = 598, Intervention II = 594; outcome 3: Total = 1,189, Intervention I = 597, Intervention II = 588). A total of 193 control group clusters (maximum 646 participants) are included in the follow-up analysis of *family* participants, which corresponds to the number of control group participants that completed the baseline questionnaire and follow-up questionnaire, excluding families where one or more members failed to complete the baseline survey. The analytic sample of *family* control group participants varied by outcome (outcome 1: control = 646; outcome 2: control = 645; outcome 3: control = 645) as detailed in Tables IV.1e and IV.1f.

Table IV.1e. Cluster and individual sample sizes by intervention status: Family Participants: Intervention I and Control Group

Number of:	Intervention sample size	Comparison sample size	Total sample size	Total response rate	Intervention response rate	Comparison response rate
Clusters						
Clusters: At beginning of study	209	215	424	N/A	NA	N/A
Clusters: Contributed at least one individual at baseline	185	197	382	90.1%	88.5%	91.6%
Clusters: Contributed at least one individual at second follow-up (timing)	179	193	372	87.7%	85.6%	89.8%
Individuals in non-attriting clusters[a]						
Individual: At time that clusters were assigned to condition	639	680	1319	N/A	NA	N/A
Individual: Contributed a baseline questionnaire	639	680	1319	100.0%	100.0%	100.0%
Individual: Contributed to second follow-up questionnaire (timing)	599	646	1245	94.4%	93.7%	95.0%
Individual: Contributed to the impact analysis at second follow-up (timing) (accounts for item non-response and any other analysis restrictions) b						
Outcome 1	598	646	1244	94.3%	93.6%	95.0%
Outcome 2	598	645	1243	94.2%	93.6%	94.9%
Outcome 3	597	645	1242	94.2%	93.4%	94.9%

Individuals were randomized after consent to participate was acquired.

Table IV.1f. Cluster and individual sample sizes by intervention status: Family Participants: Intervention II and Control Group

Number of:	Intervention sample size	Comparison sample size	Total sample size	Total response rate	Intervention response rate	Comparison response rate
Clusters						
Clusters: At beginning of study	211	215	426	N/A	N/A	N/A
Clusters: Contributed at least one individual at baseline	187	197	384	90.1%	88.6%	91.6%
Clusters: Contributed at least one individual at second follow-up (timing)	186	193	379	89.0%	88.2%	89.8%
Individuals in non-attriting clusters[a]						
Individual: At time that clusters were assigned to condition	608	680	1288	N/A	N/A	N/A
Individual: Contributed a baseline questionnaire	608	680	1288	100.0%	100.0%	100.0%

Number of:	Intervention sample size	Comparison sample size	Total sample size	Total response rate	Intervention response rate	Comparison response rate
Individual: Contributed to second follow-up questionnaire (timing)	594	646	1240	96.3%	97.7%	95.0%
Individual: Contributed to the impact analysis at second follow-up (timing) (accounts for item non-response and any other analysis restrictions) [b]						
Outcome 1	594	646	1240	96.3%	97.7%	95.0%
Outcome 2	594	645	1239	96.2%	97.7%	94.9%
Outcome 3	588	645	1233	95.7%	96.7%	94.9%

Individuals were randomized after consent to participate was acquired.

According to the cautious attrition boundary in the What Works Clearinghouse (WWC) standards, attrition was low across all study groups, participant types, and outcome scores as detailed in Tables IV.1g and IV.1h.

Table IV.1g. Rates of attrition for Intervention I compared to Control by outcome

	Intervention I		Control		Total		Attrition Classification
	N	%	N	%	N	%	N
Individual							
Randomized	649	100.0%	649	100.0%	1298	100.0%	
Attrited							
Outcome 1	181	27.9%	183	28.2%	364	28.0%	Low
Outcome 2	180	27.7%	184	28.4%	364	28.0%	Low
Outcome 3	181	27.9%	184	28.4%	365	28.1%	Low
Couple							
Randomized	404	100.0%	392	100.0%	796	100.0%	
Attrited							
Outcome 1	62	15.3%	70	17.9%	132	16.6%	Low
Outcome 2	62	15.3%	70	17.9%	132	16.6%	Low
Outcome 3	62	15.3%	71	18.1%	133	16.7%	Low
Family							
Randomized	753	100.0%	766	100.0%	1519	100.0%	
Attrited							
Outcome 1	155	20.6%	120	15.7%	275	18.1%	Low
Outcome 2	155	20.6%	121	15.8%	276	18.2%	Low
Outcome 3	156	20.7%	121	15.8%	277	18.2%	Low

Table IV.1h. Rates of attrition for Intervention II compared to Control by outcome

	Intervention II		Control		Total		Attrition Classification
	N	%	N	%	N	%	
Individual							
Randomized	654	100.0%	649	100.0%	1303	100.0%	
Attrited							
Outcome 1	184	28.1%	183	28.2%	367	28.2%	Low
Outcome 2	184	28.1%	184	28.4%	368	28.2%	Low
Outcome 3	185	28.3%	184	28.4%	369	28.3%	Low
Couple							
Randomized	392	100.0%	392	100.0%	784	100.0%	
Attrited							
Outcome 1	69	17.6%	70	17.9%	139	17.7%	Low
Outcome 2	69	17.6%	70	17.9%	139	17.7%	Low
Outcome 3	69	17.6%	71	18.1%	140	17.9%	Low
Family							
Randomized	720	100.0%	766	100.0%	1486	100.0%	
Attrited							
Outcome 1	126	17.5%	120	15.7%	246	16.6%	Low
Outcome 2	126	17.5%	121	15.8%	247	16.6%	Low
Outcome 3	132	18.3%	121	15.8%	253	17.0%	Low

B. Outcome Measures

All analyses focus on three composite outcome measures: (1) marriage and relationship skills; (2) conflict management skills; and (3) economic stability. Responses for questions within each composite measure were measured on a scale of one to four, with higher responses corresponding to more favorable outcomes. The composite score for each measure was calculated as the mean of non-missing responses. If a participant selected more than one response to a single survey question, that response was coded as missing. The Cronbach's alpha score for each outcome measure at baseline is reported to assess internal consistency.

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

Outcome name	Description of the outcome measure	Source of the measure	Timing of measure	Cronbach's Alpha
Marriage and relationship skills	The outcome measure is the average of responses to the following questions. Responses are rated on a scale of one to four, with one corresponding to a response of 'Strongly disagree' and four corresponding to a response of 'Strongly agree'. • It is important to see things from the other person's point of view • Naming feelings is necessary in good communication. • I can show others I care by my words and by my action. • I know how to let another person know that I understand their thoughts feelings and concerns. • I think about what I want to say before I speak. • I ask for what I want when something is important to me. • When family members talk together about thoughts and feelings, it is important to take turns. • It is important to show appreciation for other people's efforts. • I ask others how I can help them. • I'm respectful when others share thoughts, feelings, needs.	Local evaluation questionnaire	All survey questions are included on the pre-test, post-test (immediately after intervention ends), and the six-month follow-up test	0.87
Conflict management skills	The outcome measure is the average of responses to the following questions. Responses are rated on a scale of one to four, with one corresponding to a response of 'Strongly disagree' and four corresponding to a response of 'Strongly agree'. • I follow through on promises and agreements. • I expect others in my family to be respectful at all times. • When discussing a problem, it is important to begin by remembering the good about the other person. • The best way to solve a problem is to share ideas so that everyone is satisfied with the discussion. • I take time to relax before I bring up a problem. • A good discussion begins with setting a time and place to talk.	Local evaluation questionnaire	All survey questions are included on the pre-test, post-test (immediately after intervention ends), and the six-month follow-up test	0.82

Outcome name	Description of the outcome measure	Source of the measure	Timing of measure	Cronbach's Alpha
Economic stability	The outcome measure is the average of responses to the following questions. Responses are rated on a scale of one to four, with one corresponding to a response of 'Strongly disagree' and four corresponding to a response of 'Strongly agree'. • Learning how to make a budget is a useful way to manage money and reach goals. • Sharing feelings and concerns with family members can help lower my stress about finances. • I can reduce conflict about money through discussion and problem solving. • It is good to work from shared family values and priorities when choosing how we spend and save money. • It is important to talk about changes to traditional family roles and work through them. • I have a plan for how to save money.	Local evaluation questionnaire	All survey questions are included on the pre-test, post-test (immediately after intervention ends), and the six-month follow-up test	0.84

C. Baseline equivalence and sample characteristics

Baseline equivalence of each study group stratified by participant type (*individual*, *couple*, or *family*) was assessed by performing a chi-square test (or non-parametric test of association (i.e., Fisher's Exact Test), if appropriate due to low cell counts) of demographic attributes including participant type, gender, age group, marital status, education level, and nationality. Any attributes found to differ significantly between treatment groups were included as covariate in subsequent analyses of the outcomes. Although the RCTs are all low attrition and establishing baseline equivalence is not required, baseline outcome scores were compared by independent two-sided t-test because it is best practice.

Tables IV.3a-f include summary statistics of key baseline measures and baseline equivalence across study groups, for individual, couple and family participants completing six-month follow-up questionnaire: Intervention I and Control Group and Intervention II and Control Group. P-value of difference[1] is based on chi-square test of association if all expected cell counts were equal to or greater than five. For categorical variables with any expected cell count less than five, a Fisher's exact test was used. For outcome scores, p-values are based on a two-sided independent sample t-test.

Table IV.3a. Summary statistics of key baseline measures and baseline equivalence across study groups, for individual participants completing six-month follow-up questionnaire: Intervention I and Control Group

Baseline measure	Intervention I N (%)	Comparison N (%)	Intervention I versus comparison difference	p-value of difference[1]
Gender (%)				0.1493
Male (%)	154 (32.8%)	175 (37.6%)	-4.8%	
Female (%)	315 (67.2%)	291 (62.4%)	4.8%	
Age (%)				0.007504
<18	21 (4.5%)	18 (3.9%)	0.6%	
18-20	42 (9%)	17 (3.7%)	5.3%	
21-24	40 (8.5%)	37 (8%)	0.5%	
25-34	122 (26.1%)	126 (27.2%)	-1.1%	
35-44	132 (28.2%)	113 (24.4%)	3.8%	
45-54	62 (13.2%)	84 (18.1%)	-4.9%	
55-64	26 (5.6%)	40 (8.6%)	-3.0%	
65+	23 (4.9%)	28 (6%)	-1.1%	
Race/ethnicity (%)				0.4359
Hispanic	96 (20.5%)	118 (25.3%)	-4.8%	
Non-Hispanic Asian	106 (22.6%)	92 (19.7%)	2.9%	
Non-Hispanic Black	154 (32.8%)	138 (29.6%)	3.2%	
Non-Hispanic White	75 (16%)	74 (15.9%)	0.1%	
Non-Hispanic Other	1 (0.2%)	2 (0.4%)	-0.2%	
Unknown	37 (7.9%)	42 (9%)	-1.1%	
Marital status (%)				0.1533
Never married	127 (30.5%)	110 (25.5%)	5.0%	
Engaged	6 (1.4%)	8 (1.9%)	-0.5%	
Married	220 (52.9%)	224 (51.9%)	1.0%	
Separated	24 (5.8%)	25 (5.8%)	0.0%	
Divorced	17 (4.1%)	32 (7.4%)	-3.3%	
Widowed	22 (5.3%)	33 (7.6%)	-2.3%	
Nationality (%)				0.3703
Afghanistan	38 (8.1%)	32 (6.9%)	1.2%	
Bhutan/Nepal	65 (13.9%)	63 (13.5%)	0.4%	
Burma	30 (6.4%)	25 (5.4%)	1.0%	
Central America	45 (9.6%)	49 (10.5%)	-0.9%	
Cuba	10 (2.1%)	11 (2.4%)	-0.3%	
DRC/Burundi/Rwanda	83 (17.7%)	61 (13.1%)	4.6%	
Ethiopia/Eritrea	3 (0.6%)	1 (0.2%)	0.4%	
Iran	6 (1.3%)	2 (0.4%)	0.9%	
Iraq	31 (6.6%)	32 (6.9%)	-0.3%	
Mexico	37 (7.9%)	48 (10.3%)	-2.4%	

Baseline measure	Intervention I N (%)	Comparison N (%)	Intervention I versus comparison difference	p-value of difference[1]
Pakistan	4 (0.9%)	2 (0.4%)	0.5%	
Somalia	16 (3.4%)	30 (6.4%)	-3.0%	
Syria	7 (1.5%)	11 (2.4%)	-0.9%	
Other	93 (19.8%)	98 (21%)	-1.2%	
Unknown	1 (0.2%)	1 (0.2%)	0.0%	
Education (%)				0.152
Less than high school	178 (48%)	214 (54.2%)	-6.2%	
High school or GED	105 (28.3%)	86 (21.8%)	6.5%	
Some college, associates or vocational	41 (11.1%)	39 (9.9%)	1.2%	
Bachelors or higher	47 (12.7%)	56 (14.2%)	-1.5%	
Outcome measure 1	3.33 (0.413)	3.293 (0.418)	0.037 (0.1837)	
Outcome measure 2	3.311 (0.42)	3.276 (0.431)	0.035 (0.2138)	
Outcome measure 3	3.253 (0.449)	3.238 (0.451)	0.015 (0.6195)	
Sample size outcome 1	468	466		
Sample size outcome 2	469	465		
Sample size outcome 3	468	465		

Table IV.3b. Summary statistics of key baseline measures and baseline equivalence across study groups, for INDIVIDUAL participants completing six-month follow-up questionnaire: Intervention II and Control Group

Baseline measure	Intervention II N (%)	Comparison N (%)	Intervention II versus comparison mean difference	p-value of difference[1]
Gender (%)				0.02124
Male (%)	142 (30.2%)	175 (37.6%)	-7.4%	
Female (%)	328 (69.8%)	291 (62.4%)	7.4%	
Age (%)		(%)		0.004886
<18	16 (3.4%)	18 (3.9%)	-0.5%	
18-20	41 (8.7%)	17 (3.7%)	5.0%	
21-24	49 (10.4%)	37 (8%)	2.4%	
25-34	126 (26.8%)	126 (27.2%)	-0.4%	
35-44	123 (26.2%)	113 (24.4%)	1.8%	
45-54	64 (13.6%)	84 (18.1%)	-4.5%	
55-64	38 (8.1%)	40 (8.6%)	-0.5%	
65+	13 (2.8%)	28 (6%)	-3.2%	
Race/ethnicity (%)		(%)		0.05947
Hispanic	106 (22.6%)	118 (25.3%)	-2.7%	

Baseline measure	Intervention II N (%)	Comparison N (%)	Intervention II versus comparison mean difference	p-value of difference[1]
Non-Hispanic Asian	106 (22.6%)	92 (19.7%)	2.9%	
Non-Hispanic Black	151 (32.1%)	138 (29.6%)	2.5%	
Non-Hispanic White	84 (17.9%)	74 (15.9%)	2.0%	
Non-Hispanic Other	0 (0%)	2 (0.4%)	-0.4%	
Unknown	23 (4.9%)	42 (9%)	-4.1%	
Marital status (%)		(%)		0.1868
Never married	102 (24%)	110 (25.5%)	-1.5%	
Engaged	10 (2.4%)	8 (1.9%)	0.5%	
Married	244 (57.4%)	224 (51.9%)	5.5%	
Separated	19 (4.5%)	25 (5.8%)	-1.3%	
Divorced	16 (3.8%)	32 (7.4%)	-3.6%	
Widowed	34 (8%)	33 (7.6%)	0.4%	
Nationality (%)		(%)		0.3203
Afghanistan	37 (7.9%)	32 (6.9%)	1.0%	
Bhutan/Nepal	63 (13.4%)	63 (13.5%)	-0.1%	
Burma	30 (6.4%)	25 (5.4%)	1.0%	
Central America	48 (10.2%)	49 (10.5%)	-0.3%	
Cuba	8 (1.7%)	11 (2.4%)	-0.7%	
DRC/Burundi/Rwanda	72 (15.3%)	61 (13.1%)	2.2%	
Ethiopia/Eritrea	7 (1.5%)	1 (0.2%)	1.3%	
Iran	5 (1.1%)	2 (0.4%)	0.7%	
Iraq	29 (6.2%)	32 (6.9%)	-0.7%	
Mexico	44 (9.4%)	48 (10.3%)	-0.9%	
Pakistan	3 (0.6%)	2 (0.4%)	0.2%	
Somalia	21 (4.5%)	30 (6.4%)	-1.9%	
Syria	22 (4.7%)	11 (2.4%)	2.3%	
Other	80 (17%)	98 (21%)	-4.0%	
Unknown	1 (0.2%)	1 (0.2%)	0.0%	
Education (%)		(%)		0.9973
Less than high school	204 (54.5%)	214 (54.2%)	0.3%	
High school or GED	82 (21.9%)	86 (21.8%)	0.1%	
Some college, associates or vocational	37 (9.9%)	39 (9.9%)	0.0%	
Bachelors or higher	51 (13.6%)	56 (14.2%)	-0.6%	
Outcome measure 1	3.325 (0.394)	3.293 (0.418%)	0.032 (0.2393)	
Outcome measure 2	3.292 (0.429)	3.276 (0.431%)	0.016 (0.5772)	
Outcome measure 3	3.256 (0.43)	3.238 (0.451%)	0.018 (0.5513)	
Sample size outcome 1	470	466		
Sample size outcome 2	470	465		

Baseline measure	Intervention II N (%)	Comparison N (%)	Intervention II versus comparison mean difference	p-value of difference[1]
Sample size outcome 3	469	465		

Table IV.3c. Summary statistics of key baseline measures and baseline equivalence across study groups, for COUPLE participants completing six-month follow-up questionnaire: Intervention I and Control Group

Baseline measure	Intervention I N (%)	Comparison N (%)	Intervention I versus comparison difference	p-value of difference[1]
Gender (%)				0.8731
Male (%)	172 (50.1%)	159 (49.2%)	0.9%	
Female (%)	171 (49.9%)	164 (50.8%)	-0.9%	
Age (%)				0.7833
<18	0 (0%)	0 (0%)	0.0%	
18-20	5 (1.5%)	6 (1.9%)	-0.4%	
21-24	17 (5%)	16 (5%)	0.0%	
25-34	105 (30.6%)	95 (29.4%)	1.2%	
35-44	112 (32.7%)	100 (31%)	1.7%	
45-54	61 (17.8%)	51 (15.8%)	2.0%	
55-64	25 (7.3%)	34 (10.5%)	-3.2%	
65+	18 (5.2%)	21 (6.5%)	-1.3%	
Race/ethnicity (%)				0.6831
Hispanic	13 (3.8%)	9 (2.8%)	1.0%	
Non-Hispanic Asian	106 (30.9%)	113 (35%)	-4.1%	
Non-Hispanic Black	70 (20.4%)	58 (18%)	2.4%	
Non-Hispanic White	128 (37.3%)	125 (38.7%)	-1.4%	
Non-Hispanic Other	1 (0.3%)	1 (0.3%)	0.0%	
Unknown	25 (7.3%)	17 (5.3%)	2.0%	
Marital status (%)				0.07442
Never married	3 (0.9%)	0 (0%)	0.9%	
Engaged	0 (0%)	2 (0.6%)	-0.6%	
Married	327 (98.8%)	313 (98.4%)	0.4%	
Separated	0 (0%)	2 (0.6%)	-0.6%	
Divorced	0 (0%)	0 (0%)	0.0%	
Widowed	1 (0.3%)	1 (0.3%)	0.0%	
Nationality (%)				0.1554
Afghanistan	38 (11.1%)	40 (12.4%)	-1.3%	
Bhutan/Nepal	87 (25.4%)	95 (29.4%)	-4.0%	
Burma	14 (4.1%)	15 (4.6%)	-0.5%	

Baseline measure	Intervention I N (%)	Comparison N (%)	Intervention I versus comparison difference	p-value of difference[1]
Central America	7 (2%)	6 (1.9%)	0.1%	
Cuba	0 (0%)	1 (0.3%)	-0.3%	
DRC/Burundi/Rwanda	55 (16%)	32 (9.9%)	6.1%	
Ethiopia/Eritrea	0 (0%)	0 (0%)	0.0%	
Iran	0 (0%)	2 (0.6%)	-0.6%	
Iraq	45 (13.1%)	48 (14.9%)	-1.8%	
Mexico	5 (1.5%)	2 (0.6%)	0.9%	
Pakistan	2 (0.6%)	2 (0.6%)	0.0%	
Somalia	4 (1.2%)	2 (0.6%)	0.6%	
Syria	55 (16%)	35 (10.8%)	5.2%	
Other	29 (8.5%)	41 (12.7%)	-4.2%	
Unknown	2 (0.6%)	2 (0.6%)	0.0%	
Education (%)				0.3555
Less than high school	208 (64%)	195 (62.9%)	1.1%	
High school or GED	74 (22.8%)	61 (19.7%)	3.1%	
Some college, associates or vocational	14 (4.3%)	22 (7.1%)	-2.8%	
Bachelors or higher	29 (8.9%)	32 (10.3%)	-1.4%	
Outcome measure 1	3.343 (0.399)	3.291 (0.374)	0.052 (0.08707)	
Outcome measure 2	3.339 (0.402)	3.299 (0.394)	0.040 (0.2041)	
Outcome measure 3	3.278 (0.42)	3.219 (0.422)	0.059 (0.07969)	
Sample size outcome 1	342	322		
Sample size outcome 2	342	322		
Sample size outcome 3	342	321		

Table IV.3d. Summary statistics of key baseline measures and baseline equivalence across study groups, for COUPLE participants completing six-month follow-up questionnaire: Intervention II and Control Group

Baseline measure	Intervention II N (%)	Comparison N (%)	Intervention II versus comparison difference	p-value of difference[1]
Gender (%)				0.7529
Male (%)	164 (50.8%)	159 (49.2%)	1.6%	
Female (%)	159 (49.2%)	164 (50.8%)	-1.6%	
Age (%)				0.5332
<18	0 (0%)	0 (0%)	0.0%	
18-20	3 (0.9%)	6 (1.9%)	-1.0%	
21-24	16 (5%)	16 (5%)	0.0%	
25-34	100 (31%)	95 (29.4%)	1.6%	
35-44	94 (29.1%)	100 (31%)	-1.9%	
45-54	67 (20.7%)	51 (15.8%)	4.9%	
55-64	26 (8%)	34 (10.5%)	-2.5%	
65+	17 (5.3%)	21 (6.5%)	-1.2%	
Race/ethnicity (%)				0.952
Hispanic	11 (3.4%)	9 (2.8%)	0.6%	
Non-Hispanic Asian	112 (34.7%)	113 (35%)	-0.3%	
Non-Hispanic Black	55 (17%)	58 (18%)	-1.0%	
Non-Hispanic White	124 (38.4%)	125 (38.7%)	-0.3%	
Non-Hispanic Other	0 (0%)	1 (0.3%)	-0.3%	
Unknown	21 (6.5%)	17 (5.3%)	1.2%	
Marital status (%)				0.06211
Never married	2 (0.6%)	0 (0%)	0.6%	
Engaged	0 (0%)	2 (0.6%)	-0.6%	
Married	314 (98.7%)	313 (98.4%)	0.3%	
Separated	0 (0%)	2 (0.6%)	-0.6%	
Divorced	2 (0.6%)	0 (0%)	0.6%	
Widowed	0 (0%)	1 (0.3%)	-0.3%	
Nationality (%)				0.01249
Afghanistan	34 (10.5%)	40 (12.4%)	-1.9%	
Bhutan/Nepal	82 (25.4%)	95 (29.4%)	-4.0%	
Burma	32 (9.9%)	15 (4.6%)	5.3%	
Central America	4 (1.2%)	6 (1.9%)	-0.7%	
Cuba	2 (0.6%)	1 (0.3%)	0.3%	
DRC/Burundi/Rwanda	43 (13.3%)	32 (9.9%)	3.4%	
Ethiopia/Eritrea	0 (0%)	0 (0%)	0.0%	
Iran	0 (0%)	2 (0.6%)	-0.6%	
Iraq	37 (11.5%)	48 (14.9%)	-3.4%	
Mexico	5 (1.5%)	2 (0.6%)	0.9%	

Baseline measure	Intervention II N (%)	Comparison N (%)	Intervention II versus comparison difference	p-value of difference[1]
Pakistan	0 (0%)	2 (0.6%)	-0.6%	
Somalia	0 (0%)	2 (0.6%)	-0.6%	
Syria	56 (17.3%)	35 (10.8%)	6.5%	
Other	26 (8%)	41 (12.7%)	-4.7%	
Unknown	2 (0.6%)	2 (0.6%)	0.0%	
Education (%)				0.1396
Less than high school	203 (65.5%)	195 (62.9%)	2.6%	
High school or GED	72 (23.2%)	61 (19.7%)	3.5%	
Some college, associates or vocational	12 (3.9%)	22 (7.1%)	-3.2%	
Bachelors or higher	23 (7.4%)	32 (10.3%)	-2.9%	
Outcome measure 1	3.299 (0.404)	3.291 (0.374%)	0.008 (0.8052)	
Outcome measure 2	3.325 (0.433)	3.299 (0.394%)	0.026 (0.4331)	
Outcome measure 3	3.268 (0.432)	3.219 (0.422%)	0.049 (0.1537)	
Sample size outcome 1	323	322		
Sample size outcome 2	323	322		
Sample size outcome 3	323	321		

Table IV.3e. Summary statistics of key baseline measures and baseline equivalence across study groups, for FAMILY participants completing six-month follow-up questionnaire: Intervention I and Control Group

Baseline measure	Intervention I N (%)	Comparison N (%)	Intervention I versus comparison difference	p-value of difference[1]
Gender (%)				0.0704
Male (%)	233 (38.9%)	285 (44.1%)	-5.2%	
Female (%)	366 (61.1%)	361 (55.9%)	5.2%	
Age (%)				0.02044
<18	71 (11.9%)	57 (8.9%)	3.0%	
18-20	55 (9.2%)	80 (12.4%)	-3.2%	
21-24	65 (10.9%)	83 (12.9%)	-2.0%	
25-34	125 (20.9%)	116 (18%)	2.9%	
35-44	121 (20.3%)	100 (15.5%)	4.8%	
45-54	84 (14.1%)	97 (15.1%)	-1.0%	
55-64	44 (7.4%)	61 (9.5%)	-2.1%	
65+	32 (5.4%)	50 (7.8%)	-2.4%	
Race/ethnicity (%)				0.002499
Hispanic	77 (12.9%)	71 (11%)	1.9%	
Non-Hispanic Asian	157 (26.2%)	138 (21.4%)	4.8%	

Baseline measure	Intervention I N (%)	Comparison N (%)	Intervention I versus comparison difference	p-value of difference[1]
Non-Hispanic Black	87 (14.5%)	133 (20.6%)	-6.1%	
Non-Hispanic White	242 (40.4%)	252 (39%)	1.4%	
Non-Hispanic Other	3 (0.5%)	0 (0%)	0.5%	
Unknown	33 (5.5%)	52 (8%)	-2.5%	
Marital status (%)				0.4826
Never married	140 (27.9%)	172 (31%)	-3.1%	
Engaged	9 (1.8%)	9 (1.6%)	0.2%	
Married	288 (57.5%)	316 (57%)	0.5%	
Separated	16 (3.2%)	9 (1.6%)	1.6%	
Divorced	11 (2.2%)	8 (1.4%)	0.8%	
Widowed	37 (7.4%)	40 (7.2%)	0.2%	
Nationality (%)				0.00049
Afghanistan	91 (15.2%)	86 (13.3%)	1.9%	
Bhutan/Nepal	137 (22.9%)	124 (19.2%)	3.7%	
Burma	14 (2.3%)	20 (3.1%)	-0.8%	
Central America	43 (7.2%)	34 (5.3%)	1.9%	
Cuba	6 (1%)	7 (1.1%)	-0.1%	
DRC/Burundi/Rwanda	63 (10.5%)	101 (15.6%)	-5.1%	
Ethiopia/Eritrea	7 (1.2%)	5 (0.8%)	0.4%	
Iran	5 (0.8%)	1 (0.2%)	0.6%	
Iraq	94 (15.7%)	77 (11.9%)	3.8%	
Mexico	28 (4.7%)	15 (2.3%)	2.4%	
Pakistan	2 (0.3%)	8 (1.2%)	-0.9%	
Somalia	2 (0.3%)	3 (0.5%)	-0.2%	
Syria	72 (12%)	76 (11.8%)	0.2%	
Other	30 (5%)	89 (13.8%)	-8.8%	
Unknown	5 (0.8%)	0 (0%)	0.8%	
Education (%)				0.04696
Less than high school	310 (65.5%)	337 (66.1%)	-0.6%	
High school or GED	112 (23.7%)	107 (21%)	2.7%	
Some college, associates or vocational	19 (4%)	40 (7.8%)	-3.8%	
Bachelors or higher	32 (6.8%)	26 (5.1%)	1.7%	
Outcome measure 1	3.246 (0.365)	3.224 (0.384)	0.022 (0.3015)	
Outcome measure 2	3.236 (0.396)	3.249 (0.397)	-0.013 (0.5639)	
Outcome measure 3	3.19 (0.423)	3.177 (0.424)	0.013 (0.5945)	
Sample size outcome 1	598	646		
Sample size outcome 2	598	645		
Sample size outcome 3	597	645		

Table IV.3f. Summary statistics of key baseline measures and baseline equivalence across study groups, for FAMILY participants completing six-month follow-up questionnaire: Intervention II and Control Group

Baseline measure	Intervention II N (%)	Comparison N (%)	Intervention II versus comparison difference	p-value of difference[1]
Gender (%)				0.1173
Male (%)	235 (39.6%)	285 (44.1%)	-4.5%	
Female (%)	359 (60.4%)	361 (55.9%)	4.5%	
Age (%)				0.5087
<18	53 (9%)	57 (8.9%)	0.1%	
18-20	77 (13%)	80 (12.4%)	0.6%	
21-24	68 (11.5%)	83 (12.9%)	-1.4%	
25-34	94 (15.9%)	116 (18%)	-2.1%	
35-44	113 (19.1%)	100 (15.5%)	3.6%	
45-54	100 (16.9%)	97 (15.1%)	1.8%	
55-64	53 (9%)	61 (9.5%)	-0.5%	
65+	34 (5.7%)	50 (7.8%)	-2.1%	
Race/ethnicity (%)				0.004998
Hispanic	47 (7.9%)	71 (11%)	-3.1%	
Non-Hispanic Asian	174 (29.3%)	138 (21.4%)	7.9%	
Non-Hispanic Black	96 (16.2%)	133 (20.6%)	-4.4%	
Non-Hispanic White	243 (40.9%)	252 (39%)	1.9%	
Non-Hispanic Other	0 (0%)	0 (0%)	0.0%	
Unknown	34 (5.7%)	52 (8%)	-2.3%	
Marital status (%)				0.1095
Never married	125 (24.7%)	172 (31%)	-6.3%	
Engaged	5 (1%)	9 (1.6%)	-0.6%	
Married	321 (63.4%)	316 (57%)	6.4%	
Separated	9 (1.8%)	9 (1.6%)	0.2%	
Divorced	14 (2.8%)	8 (1.4%)	1.4%	
Widowed	32 (6.3%)	40 (7.2%)	-0.9%	
Nationality (%)				0.00049
Afghanistan	78 (13.1%)	86 (13.3%)	-0.2%	
Bhutan/Nepal	141 (23.7%)	124 (19.2%)	4.5%	
Burma	20 (3.4%)	20 (3.1%)	0.3%	
Central America	20 (3.4%)	34 (5.3%)	-1.9%	
Cuba	4 (0.7%)	7 (1.1%)	-0.4%	
DRC/Burundi/Rwanda	71 (12%)	101 (15.6%)	-3.6%	
Ethiopia/Eritrea	10 (1.7%)	5 (0.8%)	0.9%	
Iran	5 (0.8%)	1 (0.2%)	0.6%	
Iraq	68 (11.4%)	77 (11.9%)	-0.5%	
Mexico	17 (2.9%)	15 (2.3%)	0.6%	

Baseline measure	Intervention II N (%)	Comparison N (%)	Intervention II versus comparison difference	p-value of difference[1]
Pakistan	12 (2%)	8 (1.2%)	0.8%	
Somalia	0 (0%)	3 (0.5%)	-0.5%	
Syria	112 (18.9%)	76 (11.8%)	7.1%	
Other	34 (5.7%)	89 (13.8%)	-8.1%	
Unknown	2 (0.3%)	0 (0%)	0.3%	
Education (%)				0.529
Less than high school	320 (67.5%)	337 (66.1%)	1.4%	
High school or GED	102 (21.5%)	107 (21%)	0.5%	
Some college, associates or vocational	26 (5.5%)	40 (7.8%)	-2.3%	
Bachelors or higher	26 (5.5%)	26 (5.1%)	0.4%	
Outcome measure 1	3.231 (0.381)	3.224 (0.384%)	0.007 (0.7326)	
Outcome measure 2	3.209 (0.386)	3.249 (0.397%)	-0.040 (0.07288)	
Outcome measure 3	3.154 (0.421)	3.177 (0.424%)	-0.023 (0.3255)	
Sample size outcome 1	594	646		
Sample size outcome 2	594	645		
Sample size outcome 3	588	645		

Results indicated that among *individual* participants, the distribution of participants by age group was significantly different between both intervention groups and the control group. *Individual* participants in the control group were significantly more likely to be female than *individual* participants in the intervention II group as detailed in Tables IV.3a and IV.3b above.

Among *couple* participants, the distribution of marital status by study group approached, but did not meet statistical significance ($p<0.08$). Nationality was significantly different between Intervention II and control group participants as detailed in Tables IV.3c and IV3d above.

Family participants demonstrated the most demographic characteristics lacking baseline equivalence. The distribution of age group was significantly different between Intervention I and Control group participants. Race and nationality significantly varied among both intervention groups compared to the control. Details regarding the characteristics between the groups is detailed in Table IV.3e above. Additionally, the distribution of participants in each education level was significantly different among Intervention I group participants compared to the control group.

V. FINDINGS AND ESTIMATION APPROACH

A. Implementation evaluation

Key Findings:

Intervention components were implemented with fidelity and the program achieved its goals of delivering all the intended intervention components of Relationship Enhancement®, Family Stress and Conflict Management, and Financial Management. Eighty-six percent of intervention participants in the sample completed seventy-five percent or more of the program. Participants self-reported the program positively affected their relationship and communication skills. This data was collected via client interviews and is not the same as the measures developed for the impact study.

Fidelity. The RFS program was implemented by a total of 32 trained program staff across 11 program sites that received the minimum 21 hours of curriculum training. To determine fidelity, the evaluation reviewed workshop delivery in nFORM for the program. Each curriculum session was delivered in order across sites and no lesson or topic was skipped. Only those who were trained by the curriculum developer facilitated intervention sessions. There were no additional measures for fidelity constructed.

Dosage. As per attendance logged in nFORM, the data indicated that for Intervention I - Relationship Enhancement® and Family Stress and Conflict Management, seventy-two percent of the sample attended at least seventy-five percent or more of all program sessions. Ten percent of the sample did not attend any program sessions. For Intervention II - Relationship Enhancement® and Financial Management, seventy-four percent of the sample attended at least seventy-five percent or more of all program sessions. Eleven percent of the sample did not attend any program sessions. When reviewing attendance by participant type, attendance rates among *couple* and *family* participants was higher than for individual participants. Sixty-one percent of *couples* and sixty-eight percent of *family* participants completed one hundred percent or more of the program sessions, compared to fifty percent of *individual* participants. Table V.A.1. provides details of the number of participants by attendance rate.

Table V.A.1. Attendance by intervention and participant type

Participant Type	Individual	Individual	Couple	Couple	Family	Family	Total	Total
Intervention	I	II	I	II	I	II	I	II
Treatment Hours								
Mean (SD)	11.6 (6.9)	12.2 (7.0)	12.7 (5.8)	13.4 (5.4)	13.5 (5.3)	13.3 (5.6)	12.7 (6.1)	12.9 (6.1)
Median	14	16	16	16	16	16	16	16
Number of participants by attendance rate								
0%	67	71	35	30	46	64	148	165
1-24%	12	7	4	2	8	5	24	14
25-49%	37	28	20	8	21	18	78	54
50-74%	69	59	34	28	48	34	151	121
75-99%	58	59	54	46	66	68	178	173
100+	226	246	196	209	410	405	832	860

Quality. Program Officers conducted 11 site visits to observe the quality of workshop implementation and completed 49 individual participant interviews over the tenure of the local evaluation. Observations and interviews were focused on how program facilitators implemented the curriculum activities and what knowledge and skills the participants applied after completing the program. Program Officers consistently noted in each site summary report the facilitator's engagement of participants at each monitoring site visit. For each workshop observation, program officers noted the following: materials such as workbooks available in each participant's native language, interpreters were utilized when needed, and facilitators across program sites were observed consistently using white boards, PowerPoint presentations, flip charts and program materials readily available to implement activities. Feedback collected from focus group reports indicated that across all 11 sites, participants thought the facilitators were knowledgeable about the program components and were attentive to the needs of the participants.

Engagement. To measure the engagement of participants, Program Officers consistently noted in each of the site monitoring summary participant engagement by contributions to session discussions and attentiveness to the facilitators' presentations. Feedback gathered in focus groups (participant self-report) indicated that participants recommended expanding the program, so it continues for a longer length of time. Participants were able to recall key topics and skills learned in the program during both individual interviews. Individual interviews with participants consistently indicated they were applying the skills learned in the program by providing examples of their behaviors with family members and community members.

Context. To the knowledge of the RFS Program staff, no participants were enrolled in other HMRF programs. Providers and referrals available to study participants outside of the current intervention, were listed in nFORM and included county emergency and social services, child protection departments, mental health agencies, child welfare services, legal services, county and community vocational programs and educational services. The RFS Program at the YMCA of Greater Houston was paused temporarily due to Hurricane Harvey at the end of August 2017.

Programming resumed in September 2017. Although the program postponed sessions due to inclement weather, this did not have a significant impact on how the program was implemented and all programmatic activities resumed shortly afterwards. In cases of missed sessions due to weather conditions, sessions were rescheduled, or participants could attend make up sessions. There were no local policy changes that affected the implementation of HMRF programming.

Control participants did not receive any curriculum components or similar services provided by the program sites until after completing the six-month follow-up questionnaire. Control participants met with case managers as needed for linkages to providers on the resource list (e.g., county social services, child welfare services, educational services) or to collect posttest and six-month follow up data. To the knowledge of the RFS Program staff, no participants were enrolled in other HMRF programs. No formal assessment was made whether participants in the control group received similar services from another source. Informally, no instances in which this has happened are known.

B. Primary impact evaluation

Key Findings:

Three outcomes, including marriage and relationship skills, conflict management, and economic stability, were evaluated based on composite scores that had a range of possible values from one to four. Individual and couple participants in both intervention groups demonstrated higher average post-intervention scores for marriage and relationship skills, conflict management, and economic stability after controlling for the baseline scores and relevant demographics compared to the control group. Effect sizes for individual and couple participants in both intervention group comparisons were small (Cohen's D range 0.20 – 0.46). Results were more mixed for family participants, with no statistical evidence that the average post-intervention scores for any outcome were higher among family participants than the control group after controlling for baseline scores and other relevant demographics. Effect sizes for family participants in both intervention group comparisons were very small (Cohen's D < 0.15).

Outcome 1: Marriage and Relationship Skills

Among individual participants, the six-month follow-up score for outcome 1: Marriage and relationship skills, was significantly higher for participants in both intervention groups compared to the control, after controlling for baseline score. Those participants who received the intervention reported more knowledge of marriage and relationship skills at the six-month follow-up questionnaire compared to the control group, as measured by a composite outcome score with possible scores ranging from one to four. For participants in the Intervention I group who received Relationship Enhancement® plus Conflict Resolution, the mean outcome 1 score was 0.12 points higher than the control group ($p<0.001$), (Table 2.Va). This difference represented a small effect size of the intervention (0.33), (Table V4.a). For participants in the Intervention II group who received Relationship Enhancement® plus Financial Management, the mean outcome 1 score was 0.12 points higher than the control group ($p<0.001$), also corresponding to a small effect size (0.31), (Tables 2.Vb, V4.a).

Key Findings: (continued)

Among couple participants, the six-month follow-up score for outcome 1 was significantly higher for participants in both intervention groups compared to the control. For participants in the Intervention I group, the mean outcome 1 score was 0.13 points higher than the control group ($p<0.001$). For participants in the Intervention II group, the mean outcome 1 score was 0.12 points higher than the control group ($p=0.0026$). Effect sizes for couple participants in the Intervention I and Intervention II groups were higher than individual participants, however effect sizes in both groups were still considered small (0.42, 0.35, respectively).

Among *family* participants, the mean six-month follow-up score for outcome 1 was not significantly different for participants in either intervention group compared to the control at follow-up, and effect sizes were very small (<0.12 in both groups).

Outcome 2: Conflict Management Skills

The conflict management skills composite outcome score had a range of possible values from one to four. Among individual participants, the mean six-month follow-up score for outcome 2: Conflict Management Skills was significantly higher for participants in both intervention groups compared to the control after controlling for the baseline score and demographic characteristics that failed the test of baseline equivalence. For participants in the Intervention I group who received Relationship Enhancement plus Conflict Resolution, the mean outcome 2 score was 0.14 points higher than the control group ($p<0.001$), (Table V2.a). For participants in the Intervention II group who received Relationship Enhancement plus Financial Management, the mean outcome 2 score was 0.12 points higher than the control group ($p<0.001$), (Table V2.b). Effect sizes were small for both intervention groups; Intervention I had an effect size of 0.36 and Intervention II had an effect size of 0.32 (Table V4.b).

Among *couple* participants, the mean six-month follow-up score for outcome 2 was significantly higher for participants in both intervention groups compared to the control. For participants in the Intervention I group, the mean outcome 2 score was 0.15 points higher than the control group ($p=<0.001$), (Table V2.c). This observed difference represented a medium-small effect size of 0.46 (Table V4.b) For participants in the Intervention II group, the mean outcome 2 score was 0.13 points higher than the control group ($p=<0.001$), representing a small effect size of 0.37 (Tables V2.d, V4.b).

Among family participants, the mean six-month follow-up score for outcome 2 was not significantly different for participants in either intervention group compared to the control at follow-up (Tables V2.e and V2.f). Effect sizes for both intervention groups were less than 0.15 (Table V4.b).

Outcome 3: Economic Stability

The economic stability composite outcome score had a range of possible values from one to four. Among individual participants, the mean six-month follow-up score for outcome 3: Economic Stability was significantly higher for participants in both intervention groups compared to the control. For participants in the Intervention I group who received Relationship Enhancement® plus Conflict Resolution, the mean outcome 3 score was 0.09 points higher than the control group ($p=<0.001$), representing a small effect size of 0.24 (Tables V2.a, V4.c). For participants in the Intervention II group who received Relationship Enhancement plus Financial Management, the mean outcome 3 score was 0.08 points higher than the control group ($p=0.003$), representing a small effect size of 0.20 (Tables V2.b and V4.c).

Key Findings: (continued)

Among couple participants, the mean six-month follow-up score for outcome 3 was significantly higher for participants in the Intervention I group compared to the control. For participants in the Intervention I group, the mean outcome 3 score was 0.12 points higher than the control group (p=0.003), representing a small effect size of 0.36 (Tables V2.c and V4.c). For participants in the Intervention II group, the mean outcome 3 score was not statistically different than the control group (difference = 0.07, p = 0.062) and the effect size was small (0.21), (Tables V2.d and V4.c).

Among family participants, the mean six-month follow-up score for outcome 3 was not significantly different for participants in either intervention group compared to the control at follow-up (Tables V2.e and V2.f). The observed effect size for outcome 3 among family participants was very small (<0.13) for both intervention groups (Table V4.c).

For each of the three outcomes, six-month follow-up outcome scores were compared, after controlling for baseline outcome scores to determine if the treatment effects persist after program completion. Research questions were assessed by analyzing a composite of survey questions related to each outcome (See Table IV.2).

To evaluate the primary research questions, ANCOVA models were used to test for differences in six-month follow-up outcome scores across study groups, controlling for the baseline outcome score, study site, and any demographic variables that failed baseline equivalence testing (see Table V.1.a for a complete list of all variables included in each model). Separate ANCOVA models were developed for each participant type, with models for *couple* and *family* participants including a random effect to adjust for the nesting of individuals within *couple* or *family* units (See Tables V.1.b-f). The appropriateness of the ANCOVA model was determined by an examination of the following assumptions:

- Baseline outcome scores are independent of study group
- The dependent variable in the model (e.g. the post-test or follow-up outcome scores) is normally distributed
- Residuals are normally distributed and have a constant variance

Independent two-sided t-tests were used to test for differences in baseline outcome scores. Graphical and statistical tests, such as the Shapiro-Wilk test was used to examine normality. All analyses were completed using R Version 3.6.0. Findings were considered statistically significant based on a $p<0.05$ on a two-tailed test. Study group was determined based on each participant's assignment at baseline, regardless of their actual attendance of treatment sessions (i.e. analyses will adhere to intent-to-treat design). All participants who completed the baseline questionnaire and have non-missing six-month follow-up outcome scores were included in the analysis.

Table V.1.a. Covariates included in impact analyses

Model	Sample	Dependent Variable	Independent Variables	Additional Covariates
1	Intervention I Individual Participants Control Group Individual Participants	Outcome 1 Follow-up Score	Outcome 1 Baseline Score Group	Age Group Site
2	Intervention II Individual Participants Control Group Individual Participants	Outcome 1 Follow-up Score	Outcome 1 Baseline Score Group	Age Group Gender Site
3	Intervention I Couple Participants Control Group Couple Participants	Outcome 1 Follow-up Score	Outcome 1 Baseline Score Group	Marital Status Site
4	Intervention II Couple Participants Control Group Couple Participants	Outcome 1 Follow-up Score	Outcome 1 Baseline Score Group	Marital Status Nationality Site
5	Intervention I Family Participants Control Group Family Participants	Outcome 1 Follow-up Score	Outcome 1 Baseline Score Group	Race Age Group Nationality Education Site
6	Intervention II Family Participants Control Group Family Participants	Outcome 1 Follow-up Score	Outcome 1 Baseline Score Group	Race Nationality Site
7	Intervention I Individual Participants Control Group Individual Participants	Outcome 2 Follow-up Score	Outcome 2 Baseline Score Group	Age Group Site
8	Intervention II Individual Participants Control Group Individual Participants	Outcome 2 Follow-up Score	Outcome 2 Baseline Score Group	Age Group Gender Site
9	Intervention I Couple Participants Control Group Couple Participants	Outcome 2 Follow-up Score	Outcome 2 Baseline Score Group	Marital Status Site
10	Intervention II Couple Participants Control Group Couple Participants	Outcome 2 Follow-up Score	Outcome 2 Baseline Score Group	Marital Status Nationality Site
11	Intervention I Family Participants Control Group Family Participants	Outcome 2 Follow-up Score	Outcome 2 Baseline Score Group	Race Age Group Nationality Education Site

Model	Sample	Dependent Variable	Independent Variables	Additional Covariates
12	Intervention II Family Participants Control Group Family Participants	Outcome 2 Follow-up Score	Outcome 2 Baseline Score Group	Race Nationality Site
13	Intervention I Individual Participants Control Group Individual Participants	Outcome 3 Follow-up Score	Outcome 3 Baseline Score Group	Age Group Site
14	Intervention II Individual Participants Control Group Individual Participants	Outcome 3 Follow-up Score	Outcome 3 Baseline Score Group	Age Group Gender Site
Model	Sample	Dependent Variable	Independent Variables	Additional Covariates
15	Intervention I Couple Participants Control Group Couple Participants	Outcome 3 Follow-up Score	Outcome 3 Baseline Score Group	Marital Status Site
16	Intervention II Couple Participants Control Group Couple Participants	Outcome 3 Follow-up Score	Outcome 3 Baseline Score Group	Marital Status Nationality Site
17	Intervention I Family Participants Control Group Family Participants	Outcome 3 Follow-up Score	Outcome 3 Baseline Score Group	Race Age Group Nationality Education Site
18	Intervention II Family Participants Control Group Family Participants	Outcome 3 Follow-up Score	Outcome 3 Baseline Score Group	Race Nationality Site

Table V.2a. Post-intervention estimated effects using data from six-month follow-up questionnaire to address the primary research questions, Individual Participants

Outcome measure	Intervention I mean (standard deviation)	Comparison mean (standard deviation)	Intervention I: comparison mean difference (p-value of difference)
Outcome 1: Marriage and Relationship Skills	3.41 (0.39)	3.30 (0.40)	0.12 (<0.001)
Outcome 2: Conflict Management	3.41 (0.42)	3.28 (0.43)	0.14 (<0.001)
Outcome 3: Economic Stability	3.34 (0.42)	3.25 (0.44)	0.09 (<0.001
Sample Size	469	466	n/a

Source: follow-up questionnaires administered three months after baseline survey.

Notes: all models were adjusted for baseline score, age, and Site.

Table V.2b. Post-intervention estimated effects using data from six-month follow-up questionnaire to address the primary research questions, Individual Participants

Outcome measure	Intervention II mean (standard deviation)	Comparison mean (standard deviation)	Intervention II: comparison mean difference (p-value of difference)
Outcome 1: Marriage and Relationship Skills	3.42 (0.42)	3.30 (0.40)	0.12 (<0.001)
Outcome 2: Conflict Management	3.41 (0.41)	3.28 (0.43)	0.12 (<0.001)
Outcome 3: Economic Stability	3.34 (0.43)	3.25 (0.44)	0.08 (0.003)
Sample Size	470	466	n/a

Source: follow-up questionnaires administered six months after baseline survey, three months after program completion.

Notes: all models were adjusted for baseline score, age, gender and Site.

Table V.2c. Post-intervention estimated effects using data from six-month follow-up questionnaire to address the primary research questions, Couple Participants

Outcome measure	Intervention I mean (standard deviation)	Comparison mean (standard deviation)	Intervention I: comparison mean difference (p-value of difference)
Outcome 1: Marriage and Relationship Skills	3.42 (0.37)	3.30 (0.40)	0.13 (<0.001)
Outcome 2: Conflict Management	3.46 (0.40)	3.33 (0.41)	0.15 (<0.001)
Outcome 3: Economic Stability	3.38 (0.43)	3.28 (0.41)	0.12 (0.003)
Sample Size	342	322	n/a

Source: follow-up questionnaires administered six months after baseline survey, three months after program completion.

Notes: all models were adjusted for baseline score, marital status, and site.

Table V.2d. Post-intervention estimated effects using data from six-month follow-up questionnaire to address the primary research questions, Couple Participants

Outcome measure	Intervention II mean (standard deviation)	Comparison mean (standard deviation)	Intervention II: comparison mean difference (*p*-value of difference)
Outcome 1: Marriage and Relationship Skills	3.38 (0.39)	3.30 (0.40)	0.12 (0.003)
Outcome 2: Conflict Management	3.42 (0.39)	3.33 (0.41)	0.13 (0.001)
Outcome 3: Economic Stability	3.32 (0.42)	3.28 (0.41)	0.07 (0.062)
Sample Size	323	322	n/a

Source: follow-up questionnaires administered ,6 months after baseline survey, three months after program completion.

Notes: all models were adjusted for baseline score, marital status, nationality, and site.

Table V.2e. Post-intervention estimated effects using data from six-month follow-up questionnaire to address the primary research questions, Family Participants

Outcome .measure	Intervention I mean (standard deviation)	Comparison mean (standard deviation)	Intervention I: comparison mean difference (p-value of difference)
Outcome 1: Marriage and Relationship Skills	3.41 (0.39)	3.34 (0.36)	0.05 (0.1868)
Outcome 2: Conflict Management	3.42 (0.40)	3.35 (0.40)	0.06 (0.0874)
Outcome 3: Economic Stability	3.31 (0.45)	3.26 (0.42)	0.03 (0.4115)
Sample Size	598	646	n/a

Source: follow-up questionnaires administered six months after baseline, three months after program completion.

Notes: all models were adjusted for baseline score, race, age group, nationality, educational attainment, and site.

Table V.2f. Post-intervention estimated effects using data from six-month follow-up questionnaire to address the primary research questions, Family Participants

Outcome measure	Intervention II mean (standard deviation)	Comparison mean (standard deviation)	Intervention II: comparison mean difference (p-value of difference)
Outcome 1: Marriage and Relationship Skills	3.41 (0.34)	3.34 (0.36)	0.04 (0.1494)
Outcome 2: Conflict Management	3.42 (0.36)	3.35 (0.40)	0.04 (0.2607)
Outcome 3: Economic Stability	3.35 (0.40)	3.26 (0.42)	0.06 (0.0776)
Sample Size	594	646	n/a

Source: follow-up questionnaires administered six months after baseline survey, three months after program completion

Notes: all models were adjusted for baseline score, race, nationality, and site.

C. Sensitivity analyses

Key Findings:

Most (83.3%) models which included the interaction term between a binary indicator variable (completed all questions vs. completed less than all questions) and treatment group failed to demonstrate any significant difference in treatment effect between participants that completed all survey questions in the outcome scale and participants that completed fewer than all the survey questions in the outcome scale. For couple participants the mean difference in the conflict management skills score between Intervention II and Control Group participants was 0.62 points higher for participants that completed all survey questions compared to the mean difference between Intervention II and Control group participants who completed fewer than all survey questions.

Key Findings: *(continued)*

For *family* participants the mean difference in the outcome 2 score between Intervention II and Control Group participants was 0.41 points lower for participants that completed all survey questions compared to the mean difference between Intervention II and Control group participants who completed fewer than all survey questions.

For *individual* participants the mean difference in the outcome 3 score between Intervention I and Control Group participants was 0.39 points lower for participants that completed all survey questions compared to the mean difference between Intervention I and Control group participants who completed fewer than all survey questions.

Outcome scores were calculated as the mean of all non-missing responses of survey questions in the scale. Outcome scores were still calculated for participants that had missing responses on one or more survey questions in the outcome scale. Among participants in the analytic sample, 4.0% did not answer all questions in the outcome 1 scale, 2.5% did not answer all questions in the outcome 2 scale, and 3.8% did not answer all questions in the outcome 3 scale. Although the majority of participants in the analytic sample, answered all questions in the outcome scale for each of the three outcomes, inclusion of scores for participants that answered less than all survey questions may bias the results if participants were more likely to refuse questions on items they had a less favorable response on. That type of potential response bias would likely result in an overestimation of individual outcome scores.

Sensitivity analysis examined the impact of including follow-up scores for participants who did not complete all survey questions in the outcome scale on the effect of treatment results. The impact of the inclusion of outcome scores calculated on a subset of survey questions was evaluated by the inclusion of an interaction term between a binary indicator variable (completed all questions vs. completed less than all questions) and treatment group in the fully adjusted model.

Table V.3. Difference in means between intervention and comparison groups by completion of all outcome response questions methods

	Individual		Couple		Family	
	Interaction Coefficient	p-value	Interaction Coefficient	p-value	Interaction Coefficient	p-value
Outcome 1						
Intervention I	-0.11	0.4007	0.13	0.5112	0.05	0.6323
Intervention II	0.04	0.7343	0.34	0.1033	-0.06	0.4283
Outcome 2						
Intervention I	-0.26	0.2088	0.19	0.4920	-0.07	0.5938
Intervention II	-0.3	0.1365	0.62	0.0004	-0.41	<0.001
Outcome 3						
Intervention I	-0.39	0.0405	-0.12	0.6044	-0.03	0.7488
Intervention II	0.03	0.8371	0.29	0.1347	0.03	0.7356

Source: Follow-up questionnaires were administered three months after the posttest, six months after enrollment.

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

Notes: Table V.3 presents the coefficient of an interaction between study group and a binary variable indicating if the participant had completed all questions comprising the outcome measure on the six-month follow-up questionnaire. Participants who did not complete all survey question had their outcome scores calculated on a smaller number of questions than those who completed all survey questions. A significant interaction effect indicates that the effect of treatment on the follow-up score was different for Participants who completed all questions compared to those who completed fewer questions. A nonsignificant result indicates that there was no observable difference in the effect of treatment for Participants who completed all survey questions compared to those who completed fewer questions.

D. Additional analyses

Key Findings:

The observed effect size was consistently higher among individual and couple participants in both intervention groups compared to family participants across all primary outcomes. For outcome 1, individual and couple Intervention I group participants had a small-to-medium effect size (>0.30 and <0.50), however the effect size for family participants was less than 0.10 (very small). Individual and couple participants in the Intervention II group also displayed small-to-medium effect sizes, however family participants had an effect size of roughly 0.11.

For outcome 2, individual and couple Intervention I group participants had a small-to-medium effect size (>0.30 and <0.50), however the effect size for family participants was less than 0.14 (very small). Individual and couple participants in the Intervention II group also displayed small-to-medium effect sizes, however family participants had an effect size of roughly 0.09.

For outcome 3, individual and couple Intervention I group participants had a small effect size between 0.20 and 0.31), however the effect size for family participants was only 0.06. All participant types in the Intervention II group also displayed similar effect sizes between 0.10 and 0.20.

Secondary analyses address one research question 1) does the effect size for each outcome, comparing the intervention group to the control group vary by participant type (*individual*, *couple*, or *family* participant)? The secondary research question compares effect sizes from ANCOVA models of six-month follow-up outcome scores after adjusting for pre-program scores and demographic variables that failed baseline equivalence testing. Models were stratified by participant type. For ANCOVA models of couple and family participants, the model includes a random effect for each couple or family unique identifier to adjust for the nesting of individuals within couple or family groups. The secondary analysis ANCOVA models include all cases with complete data on both baseline and six-month follow-up outcome scores. The outcome scores analyzed in the secondary analysis are the same outcome scores included in the primary analysis (See Table IV.2).

To evaluate the secondary research question, the effect sizes derived from ANCOVA models were used to assess if the effect of treatment on the six-month follow-up score varied by participant type. For each outcome score, the treatment effect, comparing each intervention group to the control group were evaluated for each participant type (*individual*, *couple*, and f*amily*). The Cohen's d measure of effect size was evaluated. All analyses were completed using

R Version 3.6.0. Findings were considered statistically significant based on $p<0.05$ on a two-tailed test. No adjustment was made for clustering of participants within study sites. Study group was determined based on each participant's assignment at baseline, regardless of their actual attendance of treatment sessions (i.e. analyses adhered to an intent-to-treat design). All participants who completed the baseline questionnaire and the follow-up questionnaire and have non-missing outcome scores were included in the secondary analysis of follow-up outcomes.

The baseline outcome score was always included in ANCOVA models of the post-test and follow-up scores. Additional covariates include any demographic characteristics that failed tests of baseline equivalence (see Table V.1.a).

Table V.4a. Comparison of outcome 1 (marriage and relationship skills) post-intervention effect size (Cohen's d) among participant types using data from six-month follow-up to address the secondary research questions

Treatment Group	Individual	Couple	Family
Intervention I	0.3306	0.4183	0.1091
Intervention II	0.3068	0.3462	0.1000

Source: Second follow-up questionnaire, administered six months after the baseline questionnaire, three months after program completion.

Notes: Data presented in this table represent the effect size of each intervention compared to the control group on outcome 1 scores measured on the six-month follow-up test, adjusted for the baseline score and applicable demographic characteristics. The measure of effect shown is Cohen's' d.

Table V.4b. Comparison of outcome 2 (conflict management skills) post-intervention effect size (Cohen's d) among participant types using data from six-month follow-up to address the secondary research questions

Treatment Group	Individual	Couple	Family
Intervention I	0.3574	0.4648	0.1418
Intervention II	0.3168	0.3692	0.0782

Source: Second follow-up questionnaire, administered six months after the baseline questionnaire, three months after program completion.

Notes: Data presented in this table represent the effect size of each intervention compared to the control group on outcome 2 scores measured on the six-month follow-up test, adjusted for the baseline score and applicable demographic characteristics. The measure of effect shown is Cohen's' d.

Table V.4c. Comparison of outcome 3 (economic stability) post-intervention effect size (Cohen's d) among participant types using data from six-month follow-up to address the secondary research questions

Treatment Group	Individual	Couple	Family
Intervention I	0.2362	0.3564	0.0683
Intervention II	0.2029	0.2132	0.1235

Source: Second follow-up questionnaire, administered six months after the baseline questionnaire, three months after program completion..

Notes: Data presented in this table represent the effect size of each intervention compared to the control group on outcome 3 scores measured on the six-month follow-up test, adjusted for the baseline score and applicable demographic characteristics. The measure of effect shown is Cohen's' d.

VI. DISCUSSION

The purpose of the evaluation was to examine the effects of the RFS program. This includes the effects of the program (both Intervention I and Intervention II) on three primary program outcomes, when compared to the control group, at follow-up. The three outcomes were (1) improvement in marriage and relationship skills, (2) improvement in conflict management skills, and (3) improvement in economic stability when compared to the control group at follow-up. This included the effects of the program on these outcomes for the three different participant types: *individuals*, *couples*, and *families*.

The results were: (1) Marriage and relationships skills were higher for participants in both Intervention I and Intervention II than for participants in the control group, at follow-up. This was the case for both *individual* and *couple* participants. (2) Conflict management skills were higher for participants in both Intervention I and Intervention II than for control group participants, at follow-up. Again, this was the case for both *individual* and *couple* participants. (3) Scores for Economic stability were higher for *individual* participants in Intervention I and Intervention II, and *couple* participants in Intervention I than participants in the control group, at follow-up. *Couple* participants in Intervention II scored no differently at follow-up than control group participants. *Family* participants, in both Intervention I and Intervention II, scored no differently at follow-up than controls. This was the case for all three outcome measures.

Attendance Rates

Attendance rates among all intervention group participants varied by participant type. Among *individual* participants roughly half (48%-52%) completed 100% of the expected treatment session hours. Attendance rates among *couple* and *family* participants was higher than *individual* participants, with 67% of *couple* participants and 62% of *family* participants completing 100% or more of expected treatment session hours. Slightly more than one in ten (11.2%) of intervention group participants did not attend any treatment sessions. Attendance rates varied significantly by treatment group as indicated by a chi-square test of association Intervention II participants more likely to have higher attendance rates that Intervention I participants ($p=0.04139$).

Impact Estimation

The effect size of treatment on marriage and relationship skills and conflict management skills among *individual* and *couple* participants in both intervention groups was similar and remained in the range of small to medium effect size (range 0.29 – 0.42, mean 0.35). The effect size of treatment on economic stability was somewhat lower for *individual* participants in both intervention groups; 0.24 for participants in Intervention I and 0.20 for participants in Intervention II. For *couple* participants in Intervention II, the effect size for the economic stability score was higher than *individual* participants (0.36 compared to 0.24), but lower than the effect size related to other outcomes among couple participants in Intervention I. Although the intervention had the smallest impact on economic stability scores, economic stability may be

more likely to be influenced by external factors such as employment than the other measured outcomes.

Validity of Outcome Scores

The internal consistency for each outcome measure was examined using Cronbach's alpha. The alpha score for each outcome was examined at baseline among those participants who completed both baseline and follow-up questionnaire). The analysis revealed high internal consistency for all three outcome measures among intervention and control group participants of all three participant types (*individual, couple*, and *family*). Cronbach alpha scores ranged from 0.79 to 0.89.

Sensitivity analysis demonstrated that for the majority (83%) of models, the inclusion of outcome scores calculated on a subset of survey questions that were non-missing had no impact on the observed treatment effect. There were three models which did demonstrate that the treatment effect was significantly different among intervention participants who completed all questionnaire items compared to those who did not complete all questionnaire items. There was not, however a consistent direction of the association between treatment group and completion of all questionnaire items. Therefore, it is possible that the detected interaction effect may have been significant due to random chance.

Results in Context

Some of the study findings are straight forward. Positive changes in outcome variables can be directly attributed to the intervention. Other findings are less easily explained. One example is the study's findings regarding couple participants. Those in Intervention I scored better than controls for the marriage and relationship skills outcome and the conflict management skills outcome, but no better than controls for the economic stability outcome. To be clear, these participants scored better than controls on conflict management skills but did not take conflict management classes. They scored no better than controls on financial stability, even though they did take the financial classes. We do not have a good explanation for these findings. Perhaps the positive outcome for conflict management skills was due to "treatment group contamination;" i.e. participants in the different treatment groups spent time together outside of classes and discussed what they were learning. This is a close-knit community that may have facilitated communication among participants in the different treatment groups.

As to the economic stability factor, we do not have a good explanation for this finding either. These participants (couples in Intervention II) did take the financial management classes but did not score differently from control participants. *Couple* participants in Intervention I and *individual* participants in Intervention I did not have the financial management classes but did score significantly better than controls on the economic stability outcome. We could again argue the treatment contamination effect, because *individual* participants in Intervention II did have the financial management classes and did show improvement compared to the controls. We have no good explanation for the failure of *couple* participants in Intervention I to score better than controls.

Family participants in both Intervention I and Intervention II did not score better than controls for any of the outcome measures. Why was this? They took the classes. They should have scored better than the controls. As we were working with USCRI on the evaluation design, we indicated our support for involving participants as individuals and couples. We argued against participants enrolling as families. The response we received was that for a number of the refugee families there were several generations sharing housing arrangements. There may also be aunts, uncles, and cousins. Perhaps only one family member had a vehicle, so for multiple members of the family to participate, and they wanted to participate, they would need to come to classes at the same time. We decided that like individuals and couples, families would also be randomly assigned to group. This does not explain why the family participants assigned to the interventions scored no differently than those assigned to the control group. It may be that for many of the individuals and couples assigned to interventions, the program was seen as an opportunity to better themselves and their relationships. Perhaps for many of those in families, attendance at the classes was viewed as more of a family outing. Data providing that type of insight would be helpful, but these type data are not available.

Previous researchers have noted that interventions that featured relationship enhancement programming have been successful with Caucasian married couples, but expressed doubt about the efficacy of such programs when used with non-White, low income, and otherwise disadvantaged groups (Dion, 2005; Hawkins et al, 2008; Johnson, 2012; Karner, Bradbury, & Lavner). The results of this study, however, show that such programming can produce positive results with a refugee and immigrant population. This supports the position taken by Amato (2014), who indicated disadvantaged couples may actually be the main beneficiaries of programs of this type. It also supports Hawkins (2019) review of Healthy Marriage programming, in which he took a more positive view of program effects. Additionally, Shaw and Funk's (2019) examination of the evaluation of social service type programs provided to refugees, found that rigorous evaluation of such programs was virtually non-existent. This study used an RCT design, featuring two interventions and a control group, and did find positive results with programming targeting refugee participants.

Limitations

There are some limitations associated with the study. First, as with many studies, these are self-report data. With self-report data there are always some concerns, and thus these concerns are also applicable, but not unique to the present study. For example, we know whether or not a participant has reported an improvement in an outcome measure, but we do not have data from other sources, such family members and/or third-party observations, to corroborate what individual participants are telling us thorough their questionnaire responses. Improvements shown may be due to participants responding based on perceived social desirability, rather than actual improvement.

Study results for among those participating as individuals or in couples showed a number of positive benefits three months after completion of the program. This is a relatively short follow-up period. Did the changes we observed at follow-up continue? If so, did that improvement

translate into relationship, social, and financial success? Data to answer these questions are not available, but extending the follow-up associated with the evaluation of this type of program, and exploring the degree to which the type of positive results seen in this study translate into real world success should be subjects of future research.

Additionally, participants represented at least 15 different language groups. Even when participants spoke a language fluently, level of literacy varied greatly. Thus, there may be some concerns as to whether all participants clearly understood all of the questionnaire items to which they responded. To address this concern, extensive efforts were made to provide each participant questionnaires in his/her own language and to have interpreters available to aid participants in completing questionnaires, as needed. Every study has its limitations, and despite the limitations identified here, we believe these results are quite encouraging, providing important information about the positive effects of the RFS program and its potential to provide real assistance to refugees who are resettling in the U.S.

REFERENCES

Accordino, M. P., & Guerney, B. G. (2002). *The empirical validation of relationship enhancement couple and family therapy*. In D. J. Cain (Ed.), *Humanistic psychotherapies: Handbook of research and practice* (p. 403–442). American Psychological Association.

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(3), 343–355.

Bellg, A. J., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D. S., Ory, M., Czajkowski, S. (2004). Enhancing Treatment Fidelity in Health Behavior Change Studies: Best Practices and Recommendations from the NIH Behavior Change Consortium. *Health Psychology*, 23(5), 443–451.

Bellinger, G. A. (2013). Negotiation of gender responsibilities in resettled refugee populations through Relationship Enhancement training. *Transcultural Psychiatry*, 50(3), 455–471.

Carey, R. (2016). Refugee Resettlement in the United States. (2016, March 21). Administration for Children & Families Archives. Retrieved August 19, 2019, from https://www.acf.hhs.gov/archive/blog/2016/03/refugee-resettlement-in-the-united-states

Conner, P. Krogstad, J. M. (2018). Many worldwide oppose more migration-both into and out of their countries. Fact Tank, News in The Numbers, December 10, 2018. Retrieved August 8, 2020 from https://www.pewresearch.org/fact-tank/2018/12/10/many-worldwide-oppose-more-migration-both-into-and-out-of-their-countries/

Critical Appraisal Skills Programme. (2018) CASP checklists: Critical appraisal skills programme, Oxford, UK. Retrieved from https://casp-uk.net/casp-tools-checklists/.

Dion, M. R. (2005). Healthy Marriage Programs: Learning What Works. *The Future of Children*, 15(2), 139–156.

Guerney, B. G. (1977) Relationship Enhancement, Jossey-Bass, San Francisco

Guerney, B. G., Ortwein, M., & Amin, G. (2007). Relationship Enhancement for refugees and immigrants. Leader's guide. Frankfort, KY: Institute for Development of Emotional and Life Skills and U.S. Committee for Refugees and Immigrants.

Guerney, B. G., Ortwein, M., Amin, G., Chaterdon, M., & Solly, A. (2008). Relationship enhancement and financial literacy. Leader's guide. Washington, DC: U.S. Committee for Refugees and Immigrants.

Hawkins, A. J. (2019). Are federally supported relationship education programs for lower-income individuals and couples working? A review of evaluation research. American Enterprise Institute. Retrieved, July 16, 2020 from https://www.aei.org/wp-content/uploads/2019/08/Are-Federally-Supported-Relationship-Education-Programs-for-Lower-Income-Individuals-and-Couples-Working.pdf

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(5), 723–734.

Jackson, G. L., Trail, T. E., Kennedy, D. P., Williamson, H. C., Bradbury, T. N., & Karney, B. R. (2016). The salience and severity of relationship problems among low-income couples. *Journal of Family Psychology*, 30(1), 2–11. Johnson, M. D. (2012). Healthy marriage initiatives: On the need for empiricism in policy implementation. American Psychologist, 67(4), 296-308.

Karney, B. R., Bradbury, T. N., & Lavner, J. A. (2018). Supporting healthy relationships in low-income couples: Lessons learned and policy implications. *Policy Insights from the Behavioral and Brain Sciences*, 5(1), 33-39.

Krogstad, J. M. (2019, October 7). Key facts about refugees to the U.S. Retrieved November 5, 2019, from https://www.pewresearch.org/fact-tank/2019/10/07/key-facts-about-refugees-to-the-u-s/.

Krogstad, J.M & Igielnik, R. (2017, February 3). Where refugees to the US come from. Retrieved August 20, 2019, from https://www.pewresearch.org/fact-tank/2017/02/03/where-refugees-to-the-u-s-come-from/.

Mansson, J., & Delander, L. (2017). Mentoring as a way of integrating refugees into the labour market—Evidence from a Swedish pilot scheme. *Economic Analysis and Policy*, 56, 51–59.

Office of Refugee Resettlement (2014) Annual Report to Congress FY 2014. Retrieved from https://www.acf.hhs.gov/sites/default/files/orr/orr_annual_report_to_congress_fy_2014_ signed.pdf

Office of Refugee Resettlement (2015). The U.S. Refugee Resettlement Program – an Overview. (2015, September 14). Retrieved November 5, 2019, from https://www.acf.hhs.gov/orr/resource/the-us-refugee-resettlement-program-an-overview.

Ortwein, M. C., & Amin, G. (2011). Relationship Enhancement for refugees and immigrants 2nd edition. Illustrated participant manual [Translated in Arabic, Burmese, Farsi, Karen, Nepali, Somali, and Tigrinya]. Arlington, VA: U.S. Committee for Refugees and Immigrants.

Ortwein, M. (2014). Relationship Enhancement for refugees and immigrants. Leader's guide. Frankfort, KY: Institute for Development of Emotional and Life Skills and U.S. Committee for Refugees and Immigrants.

Radford, J. & Connor, P. (2019). These are countries that accept the most refugees in the world. World Economic Forum. https://www.weforum.org/agenda/2019/06/canada-now-leads-the-world-in-refugee-resettlement-surpassing-the-u-s/

Shaw, S. A., & Funk, M. (2019). A Systematic Review of Social Service Programs Serving Refugees. *Research on Social Work Practice*, 29(8), 847–862.

Stark, L., Seff, I., Assezenew, A., Eoomkham, J., Falb, K., & Ssewamala, F. M. (2018). Effects of a Social Empowerment Intervention on Economic Vulnerability for Adolescent Refugee Girls in Ethiopia. *Journal of Adolescent Health*, 62(1), S15-S20.

Uehling, G. (2015). The Responsibilization of Refugees in the United States: On the Political Uses of Psychology. *Anthropological Quarterly*, 88(4), 997-1028.

United Nations High Commissioner for Refugees. (n.d.a) Retrieved from https://www.unhcr.org/en-us/history-of-unhcr.html, accessed, November 4, 2019.

United Nations High Commissioner for Refugees. (n.d.b). Figures at a glance. Retrieved from http://www.unhcr.org/figures-at-a-glance.html

United Nations High Commissioner for Refugees. (2018). Resettlement at a glance: 2017 in review. Retrieved from http://www.unhcr.org/en-us/5a9d507f7

Urban Institute (n.d.). Private-Sector Partners for Supporting Refugees. Retrieved from https://apps.urban.org/features/refugee-support-partnerships/

Williamson, H. C., Altman, N., Hsueh, J., & Bradbury, T. N. (2016). Effects of relationship education on couple communication and satisfaction: A randomized controlled trial with low-income couples. *Journal of Consulting and Clinical Psychology*, 84(2), 156–166.

VIII. APPENDICES

A. Logic Model (or Theory of Change) for the program

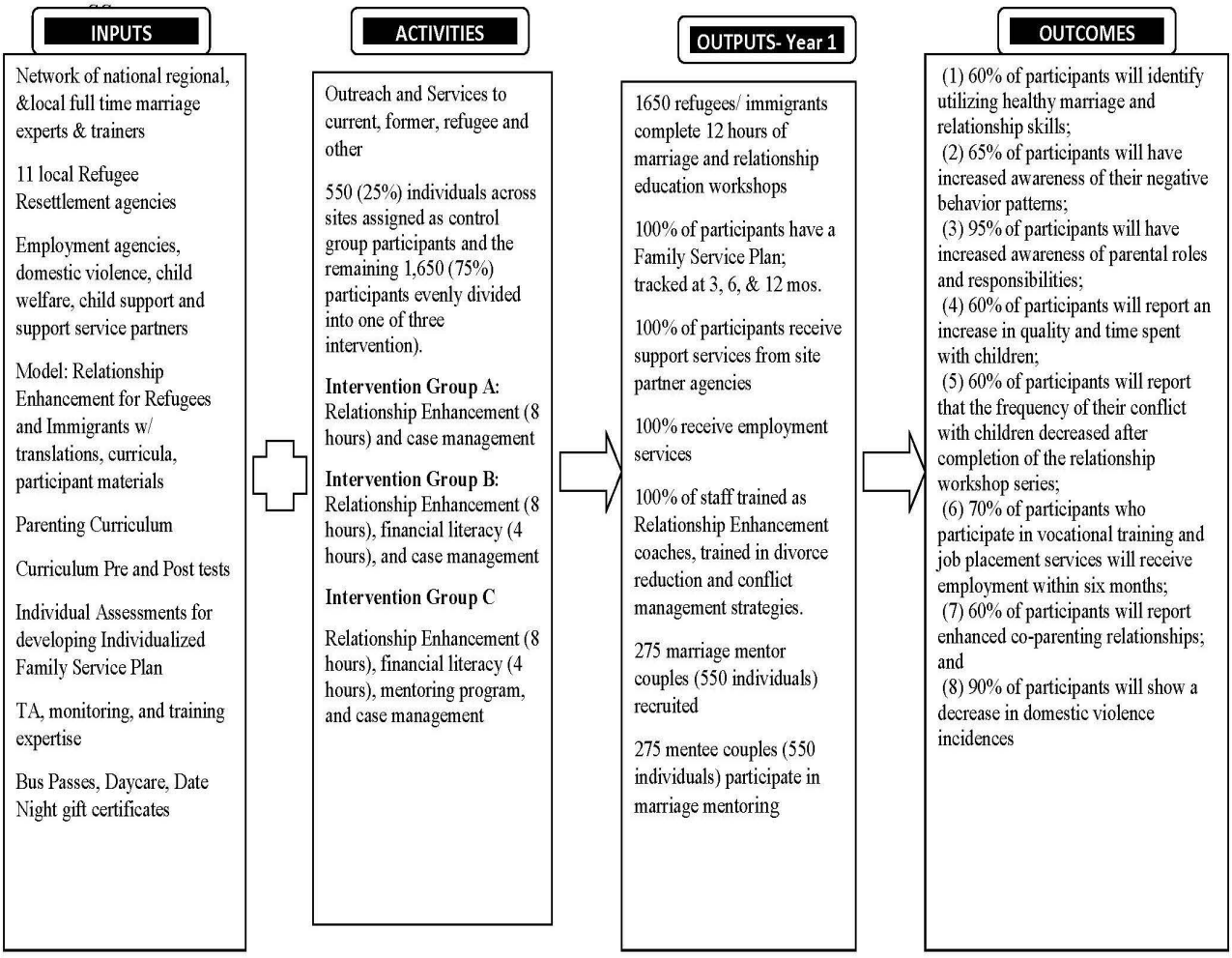
B. Data and Study Sample

C. Attrition Rates and Baseline Equivalence of the RCT Design.

D. Additional Analyses

Appendix A: Logic Model (or Theory of Change) for the program

Figure A.1. Logical Model



Appendix B: Data, Sample, and Measures

Table B.1. Data used to address implementation research questions – Intervention I

Implementation element	Research question	Data source	Timing/frequency of data collection	Party responsible for data collection
Fidelity	Were all intended intervention components of Relationship Enhancement® and Family Stress and Conflict Management delivered as proposed during curriculum training?	Workshop sessions in nFORM; Operational reports in nFORM	All sessions delivered	Program Coordinators Evaluation Staff
Fidelity	Who facilitated each of the RFS workshop sessions across the 11 sites?	Attendance logs; session observations; sessions logged in nFORM; Operational Reports in nFORM – Session Attendance	nFORM session attendance logged weekly; bi-weekly staff meeting; annual site visits (2 to 3 sites selected each year)	Program Coordinators, Program Officers
Dosage	What was the average amount of curriculum hours that participants received?	Workshop sessions and individual service contacts in nFORM; nFORM Data Export File	All sessions delivered	Program Coordinators, Case Managers
Quality	What was the quality of staff–participant interactions?	Observations of interaction quality during workshop facilitation monitoring; individual interviews with participants	annual monitoring site visits (2 to 3 sites selected each year)	Program Officers
Engagement	How engaged were intervention group members in the program?	Observations of interaction quality during workshop facilitation monitoring; focus group interviews	annual monitoring site visits, focus group interviews (2 to 3 sites selected each year)	Local Evaluation Staff Program Officers
Context	What, if any, additional programming and resources were made available to participants?	Interviews with Program Coordinators and Case Managers, list of activities/resources available at each participating site and listed in nFORM under Provider	Bi-weekly Meetings, monthly check-ins; ad hoc	Program Coordinators, Program Officers
Context	What external events affected implementation?	Yearly meeting with USCRI Staff and RFS Program staff; list of services provided by partner agencies/sites	Once a year; ad hoc	Program Coordinators, Case Managers, Program Officers

Table B.2. Data used to address implementation research questions – Intervention II

Implementation element	Research question	Data source	Timing/frequency of data collection	Party responsible for data collection
Fidelity	Were all intended intervention components of Relationship Enhancement® and Financial Management delivered as proposed during curriculum training?	Workshop sessions in nFORM; Operational reports in nFORM	All sessions delivered	Program Coordinators Evaluation Staff
Fidelity	Who facilitated each of the RFS workshop sessions across the 11 sites?	Attendance logs; session observations; sessions logged in nFORM; Operational Reports in nFORM – Session Attendance	nFORM session attendance logged weekly; bi-weekly staff meeting	Program Coordinators, Program Officers
Dosage	What was the average amount of curriculum hours that participants received?	Workshop sessions and individual service contacts in nFORM; nFORM Data Export File	All sessions delivered	Program Coordinators, Case Managers
Quality	What was the quality of staff–participant interactions?	Observations of interaction quality during workshop facilitation monitoring; individual interviews with participants	Annual monitoring site visits (2 to 3 sites selected each year)	Program Officers
Engagement	How engaged were intervention group members in the program?	Observations of interaction quality during workshop facilitation monitoring; with participants; focus group interviews	annual monitoring site visits, focus group interviews (2 to 3 sites selected each year)	
Context	What, if any, additional programming and resources were made available to participants?	Interviews with Program Coordinators and Case Managers, list of activities/resources available at each participating site and listed in nFORM under provider	Bi-weekly Meetings, monthly check-ins; ad hoc	Program Coordinators, Program Officers
Context	What external events affected implementation?	Yearly meetings with USCRI Staff and RFS Program staff; list of services provided by partner agencies/sites	Once a year; ad hoc	Program Coordinators, Case Managers, Program Officers

Table B.3. Data used to address implementation research questions – Counterfactual

Implementation element	Research question	Data source	Timing/frequency of data collection	Party responsible for data collection
Fidelity	What content did the control group receive from any sources during the evaluation period?	Case management notes in nFORM if applicable	Monthly	Program Coordinators
Dosage	How often did the control group members participate in services similar to the intervention from other sources, on average?	Case management notes in nFORM if applicable	Monthly	Program Coordinators
Quality	What was the quality of staff–participant interactions in services similar to the intervention received from other sources, if known?	Observations of interaction quality during site visits	annual site visits (2 to 3 sites selected each year)	Program Officers
Engagement	How engaged were control group members in services similar to the intervention received from other sources, if known?	Case management notes in nFORM if applicable	Monthly	Program Coordinators, Program Officers
Context	What, if any, programming and resources were made available to participants?	Interviews with Program Coordinators and Case Managers, list of activities/resources available at each participating site	Bi-weekly Meetings; ad hoc	Program Coordinators, Program Officers

The nFORM dataset was exported and organized into separate attendance categories to determine the number of hours completed by participant type and intervention group. nFORM operational reports indicating workshop dates for each session was reviewed to check facilitator logged to ensure that all sessions were delivered by a facilitator who completed the required curriculum training. Operational reports provided class dates and hours delivered for sessions. The most frequently reported experiences in focus group reports were reviewed to and reported to measure engagement. Summative monitoring reports for each site included information gathered through workshop observations, staff, and client interviews. The summaries were reviewed to determine if the following criteria were met during observations to measure quality and engagement: verbal interactions with participants including discussions, use of instruction materials and technology during workshop, availability of materials to participants, and participant self-report of a skill learned in the program. To determine if an external event affected implementation, reports made during bi-weekly team meetings and monthly check-ins with each site were reviewed.

Table B.4. Key features of the impact analysis data collection

	Data source	Timing of data collection	Mode of data collection	Party responsible for data collection	Start and end date of data collection
Intervention I	Intervention group study participants	Enrollment (baseline)	In-person online survey (nFORM) In-person online survey (Qualtrics)	Program Coordinator, Case Manager, Data Collection Specialist	July 1, 2016 through September 29, 2019
		Last day of session/end of curriculum classes (posttest)	In-person online survey (nFORM) In-person online survey (Qualtrics)	Program Coordinator, Case Manager, Data Collection Specialist	October 1, 2016 through December 31, 2019
		Six months following baseline (follow-up)	In-person online survey or via phone if person is unable to come in (Qualtrics)	Program Coordinator, Case Manager, Data Collection Specialist	December 1, 2016 through March 31, 2020
Intervention II	Intervention group study participants	Enrollment (baseline)	In-person online survey (nFORM) In-person online survey (Qualtrics)	Program Coordinator, Case Manager, Data Collection Specialist	July 1, 2016 through September 29, 2019
		Last day of session/end of curriculum classes (posttest)	In-person online survey (nFORM) In-person online survey (Qualtrics)	Program Coordinator, Case Manager, Data Collection Specialist	October 1, 2016 through December 31, 2019
		Six months following baseline (follow-up)	In-person online survey or via phone if person is unable to come in (Qualtrics)	Program Coordinator, Case Manager, Data Collection Specialist	December 1, 2016 through March 31, 2020
Counterfactual	Comparison group study participants	Enrollment (baseline)	In-person online survey (nFORM) In-person online survey (Qualtrics)	Program Coordinator, Case Manager, Data Collection Specialist	July 1, 2016 through September 29, 2019
		Last day of session/end of curriculum classes (posttest)	In-person online survey (nFORM) In-person online survey (Qualtrics)	Program Coordinator, Case Manager, Data Collection Specialist	October 1, 2016 through December 31, 2019
		Six months following baseline (follow-up)	In-person online survey or via phone if person is unable to come in (Qualtrics)	Program Coordinator, Case Manager, Data Collection Specialist	December 1, 2016 through March 31, 2020

Appendix C: Randomization assignment by site and participant type

The intended probability was for each *individual* participant, *couple*, or *family* to have a one in three probability of assignment to a specific group. There was no additional sub-sampling that occurred after random assignment, chi-square test of group by site in the analytic sample found no evidence of an association.

Table C.I. Randomization Assignment by site and participant type

Site	Participant Type	Intervention I	Intervention II	Control	p-value
Albany	Individual	28	26	26	0.9976
Colchester	Individual	36	41	41	
Des Moines	Individual	9	7	9	
Detroit	Individual	16	15	15	
Erie	Individual	17	18	17	
Houston	Individual	51	49	48	
Kansas City	Individual	135	122	130	
Los Angeles	Individual	80	87	90	
Philadelphia	Individual	86	80	84	
Raleigh	Individual	33	32	28	
St. Louis	Individual	18	24	21	
Site	**Participant Type**	**Intervention I**	**Intervention II**	**Control**	**p-value**
Albany	Couple	30	30	32	0.9839
Colchester	Couple	33	28	39	
Des Moines	Couple	20	26	25	
Detroit	Couple	32	24	32	
Erie	Couple	72	68	68	
Houston	Couple	46	56	46	
Kansas City	Couple	60	54	48	
Los Angeles	Couple	10	6	8	
Philadelphia	Couple	22	24	22	
Raleigh	Couple	20	16	18	
St. Louis	Couple	16	13	12	
Site	**Participant Type**	**Intervention I**	**Intervention II**	**Control**	**p-value**
Albany	Family	71	84	81	0.3306
Colchester	Family	22	16	29	
Des Moines	Family	88	90	106	
Detroit	Family	99	88	90	
Erie	Family	97	109	105	
Houston	Family	61	72	67	
Kansas City	Family	13	6	12	
Los Angeles	Family	67	38	61	

Site	Participant Type	Intervention I	Intervention II	Control	p-value
Philadelphia	Family	33	36	39	
Raleigh	Family	71	51	73	
St. Louis	Family	17	18	17	

Appendix D: Data preparation

The local evaluation questionnaire and nFORM datasets were merged based on participant ID and survey type (baseline, post-program, follow-up). The combined dataset underwent an extensive review and cleaning process each quarter. Data cleaning included identifying and resolving duplicates, coding survey questions as missing if a participant selected more than one response, identifying inconsistent values across data collection instruments and collection time points, identifying implausible values, identifying missing values, and identifying participants with missing questionnaires and attendance data, if applicable. A data quality report was generated each quarter to help program coordinators correct data or to collect missing data as needed. For example, each quarter data was assessed to ensure that each participant who was categorized as a couple or family participant type had one or more additional participants with a matching couple or family identification number. Additional data quality checks included identifying participants without a recorded treatment group, participant type, family identification number, gender, or nationality, participants assigned to more than one study site, participants who have completed a six-month evaluation questionnaire but not a post-program questionnaire, and participants with inconsistent demographic data between the nFORM and local evaluation questionnaire. In cases where inconsistent or implausible values could not be corrected, responses were coded as 'Unknown'.

After attempts to correct data and to collect missing information were completed, missingness was evaluated by tabulating the number of missing values by treatment group, study site, and demographic characteristics. Differential probability of missingness of the dependent variable was accounted for in the analysis by including only cases with both baseline and six-month follow-up scores and any covariate(s) associated with missingness in ANCOVA models. Differential probability of missingness in any relevant baseline variables such as demographic characteristics was accounted for by including a 'Missing/Unknown' category as a classification variable in ANCOVA models.

Figure D.1. CONSORT diagram for individual Participants, for studies in which consent occurred before assignment

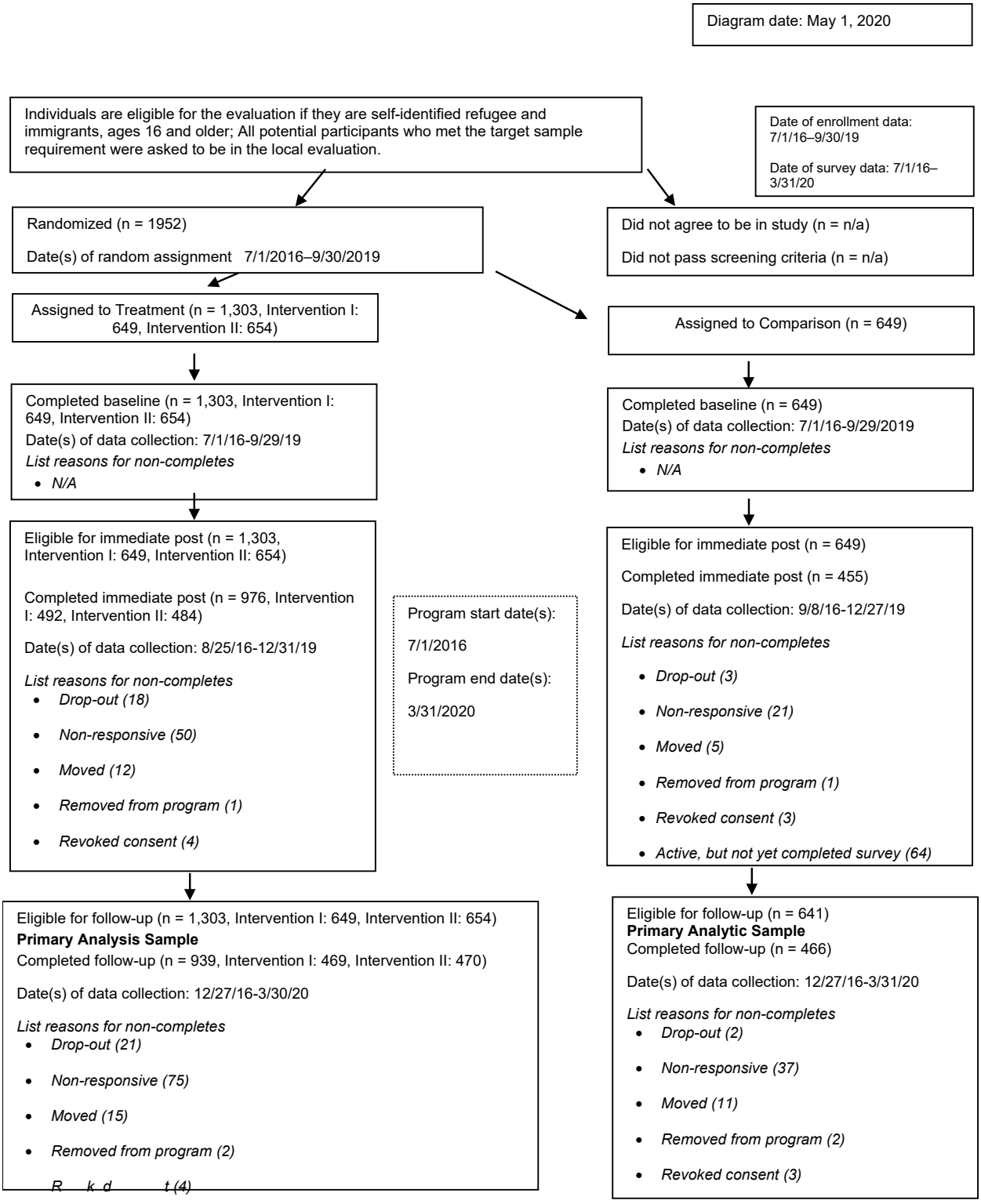


Figure D.2. CONSORT diagram for couple clusters, for studies in which consent occurred before assignment

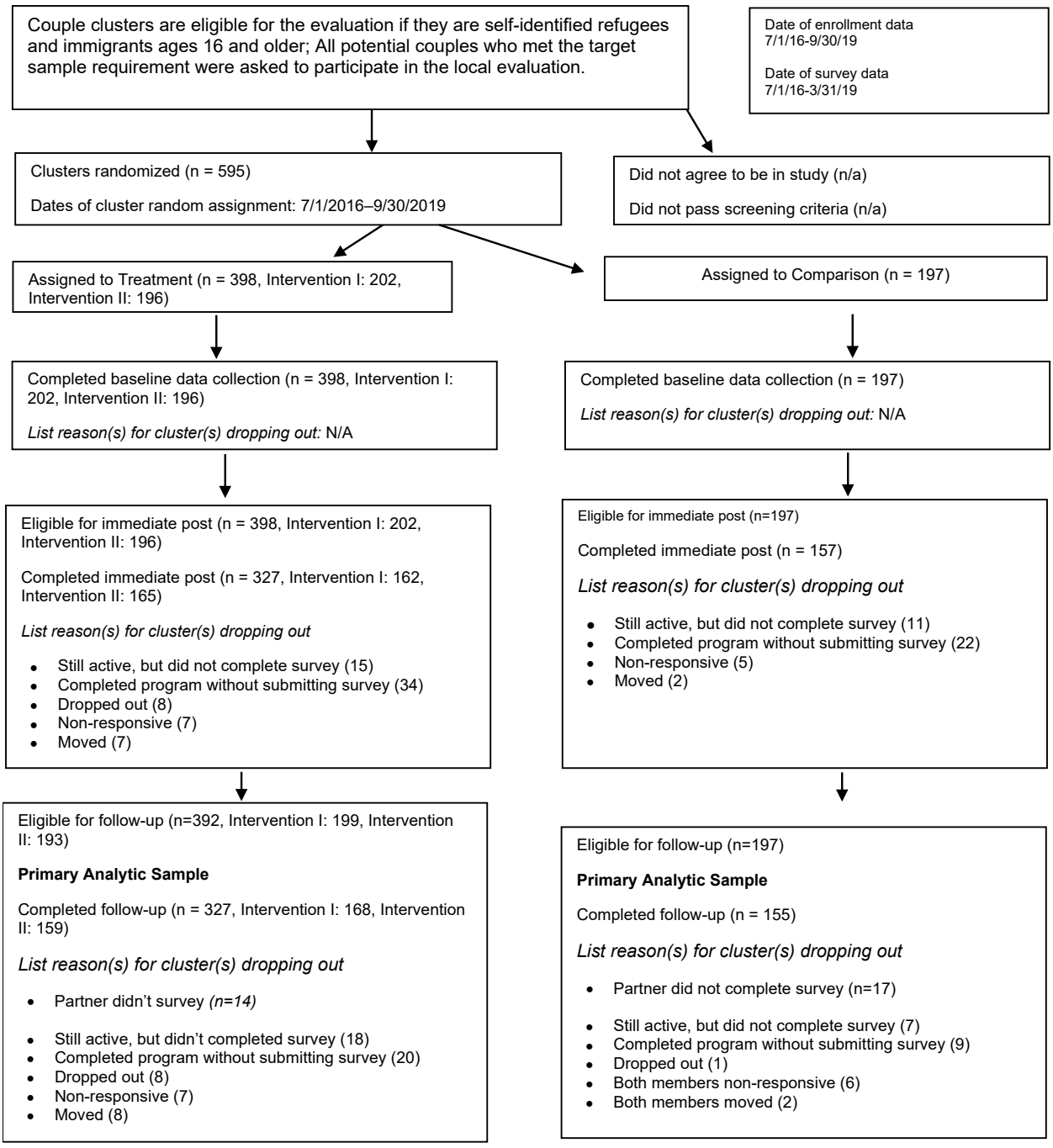


Figure D.3. CONSORT diagram for individual Participants within couples for studies in which consent occurred before assignment

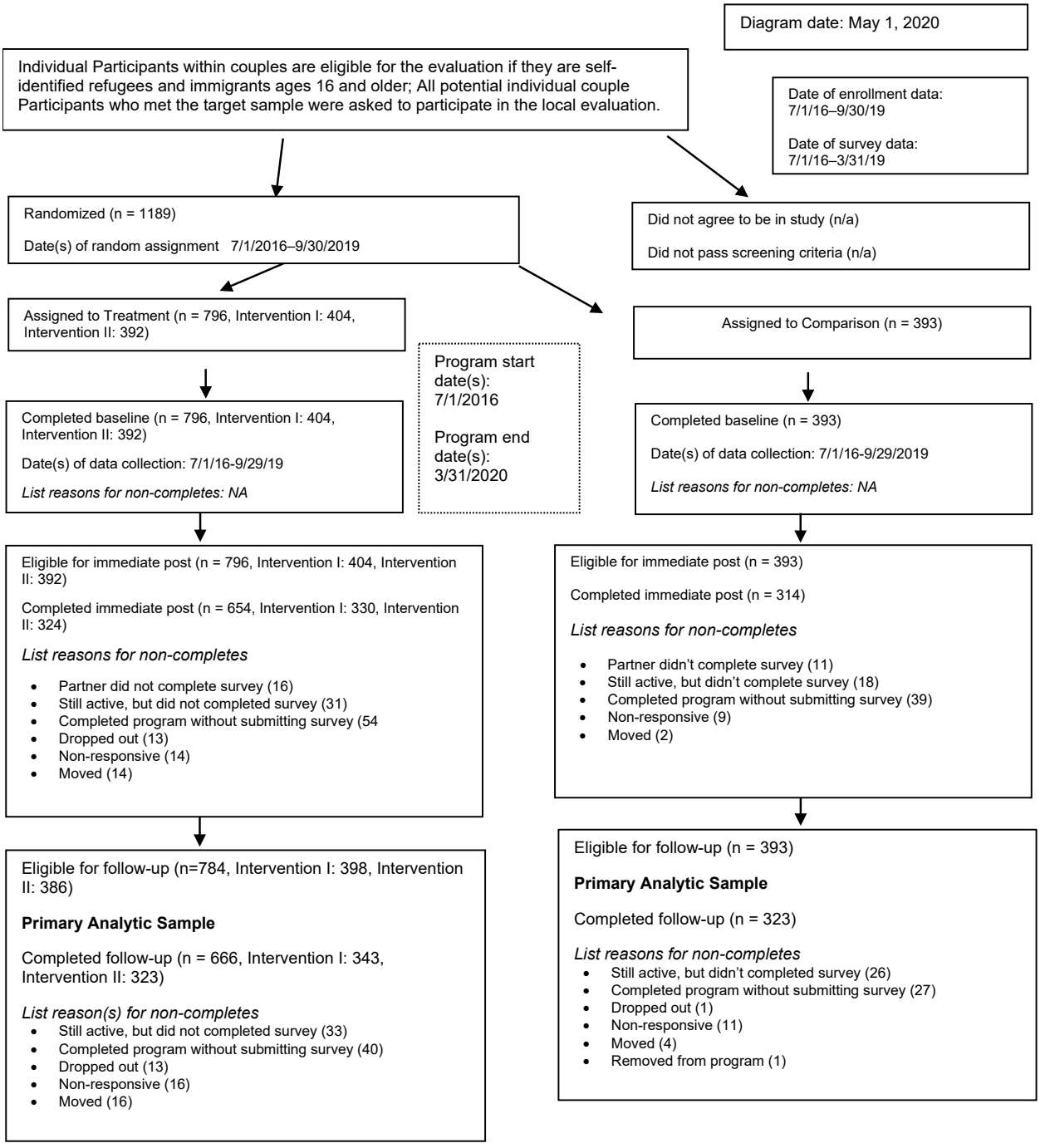


Figure D.4. CONSORT diagram for family clusters, for studies in which consent occurred <u>before</u> assignment

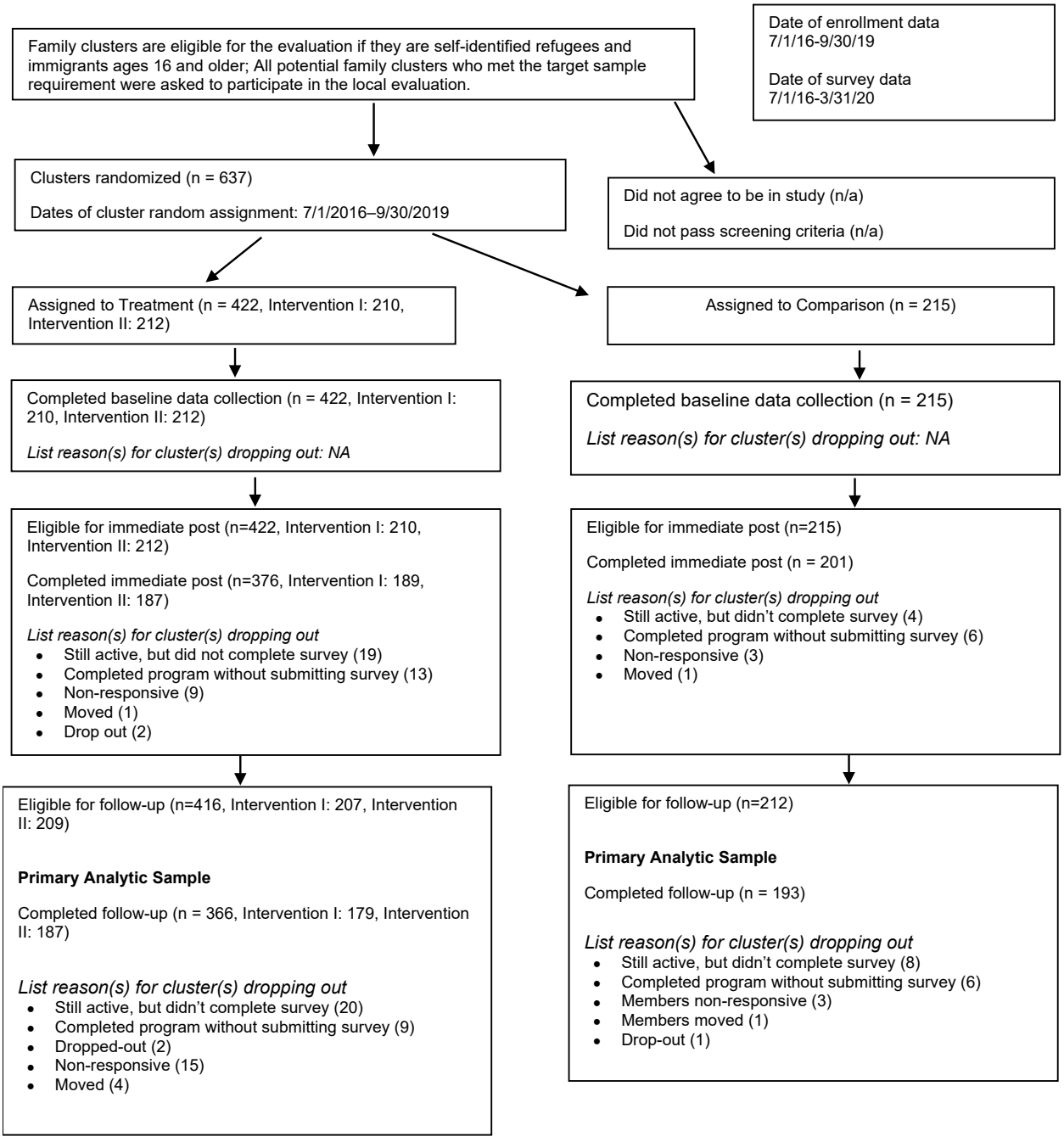


Figure D.5. CONSORT diagram for individual Participants within families, for studies in which consent occurred <u>before</u> assignment