



Evaluation of OurRelationship and ePREP Online Couple Relationship
Education Programs

Final Impact Evaluation Report for
the Administration for Children and Families

October 28, 2020

Prepared by

Galena K. Rhoades, PhD

Maggie O'Reilly Treter, MA

Kayla Knopp, PhD

University of Denver

Recommended Citation:

Rhoades, G. K., O'Reilly Treter, M., & Knopp, K., (2020). Impact Evaluation of OurRelationship and ePREP Online Couple Relationship Education Programs. University of Denver.

Acknowledgements:

We would like to thank the Administration for Children and Family's Office of Planning, Research, and Evaluation (OPRE) for their support of this project, as well as Mathematica Policy Research for their guidance and consultation. We would also like to thank the staff at the University of Miami and the University of Denver for their collaboration in managing and executing the project. Finally, we would like to thank the ePREP and OurRelationship Coaches for their work in supporting program participants.

Disclosure:

Dr. Brian Doss, the Project Director, is a coinventor of the intellectual property used in this study and an equity owner in OurRelationship LCC.

This publication was prepared under Grant Number 90FM0063 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: “Evaluation of OurRelationship and ePREP Online Couple Relationship Education Programs”

Objective. Low-income individuals experience especially high rates of relationship distress and/or divorce (Bramlett & Mosher, 2002; Lundquist et al., 2013; Jackson, Krull, Bradbury, & Karney, 2017). However, there are several barriers to low-income couples receiving relationship education. The current study investigated the effectiveness of two web-based interventions for low-income couples.

Study design. This impact evaluation was conducted in two phases. Phase I included 742 low-income couples who were recruited nationally and randomized to the online OurRelationship program (Doss et al., 2016), the online ePREP program (Braithwaite & Fincham, 2011), both with coaching, or a Waitlist Control group. Phase II included 1,250 low-income couples who were recruited nationally and randomized to the OurRelationship program or the ePREP program with or without coaching or a Waitlist Control group. If a couple was assigned to a No Coach condition and did not engage well with the program, they were re-assigned to receive a Contingent Coach or not (i.e., remaining in the No Coach condition). In both Phase I and Phase II, couples completed the four-item version of the Couple Satisfaction Index (CSI-4; Funk & Rogge, 2007) at enrollment and six months later. For both phases, the final sample analyzed in the current report included couples wherein both partners completed the CSI at baseline and the six-month follow-up (Phase I $N=616$ couples, Phase II $N=747$ couples as of May 1, 2020).

Results. In Phase I, couples assigned to ePREP and OurRelationship experienced significantly higher relationship satisfaction six months after randomization compared to the Waitlist Control group. The two interventions did not significantly differ from one another in terms of relationship satisfaction. In Phase II, couples assigned to Coach, Contingent Coach, or No Coach did not report significantly different levels of relationship satisfaction from one another six months after randomization, but they did report higher relationship satisfaction compared to the Waitlist Control group.

Conclusion. Results indicate that brief, web-based interventions can serve an important role in delivering effective services to low-income couples. Couples who received coach contact were more likely to complete the program compared to couples who had limited or no coach contact, but providing coach contact did not improve relationship satisfaction above and beyond the online programs themselves six months after randomization.

Contents

I.	Introduction	1
A.	Introduction and study overview	1
B.	Primary research question(s)	2
C.	Secondary research question(s).....	3
II.	Intervention and counterfactual conditions	4
A.	Description of program as intended.....	4
B.	Description of counterfactual condition as intended.....	8
C.	Research Questions about the intervention and counterfactual conditions as implemented	9
III.	Study design	10
A.	Sample formation and research design	10
B.	Data collection	14
IV.	Analysis methods	15
A.	Analytic sample	15
B.	Outcome measures	19
C.	Baseline equivalence and sample characteristics.....	19
V.	Findings and Estimation Approach	24
A.	Implementation evaluation	24
B.	Primary impact evaluation	26
VI.	Discussion	36
VII.	References	39
VIII.	Appendices.....	40
A.	Data and study sample.....	40
B.	Baseline equivalence summary findings and matching approach in Phase II	44
C.	DATA PREPARATION	52
D.	IMPACT ESTIMATION.....	52
E.	ADDITIONAL TABLES	53

Tables

II.1.	Description of intended intervention and counterfactual components and target populations for Phase I and Phase II	7
II.2.	Staff training and development to support intervention and counterfactual components.....	8
II.3.	Research questions about implementation of OurRelationship/ePREP and the Waitlist Control Group	9
IV.1a.	Phase I - Individual sample sizes for the analytic sample by intervention status	16
IV.1b.	Phase II - Individual sample sizes for the analytic sample by intervention status – OurRelationship.....	17
IV.1c.	Phase II - Individual sample sizes for the analytic sample by intervention status – ePREP	18
IV.2.	Outcome measure used for primary impact analyses research questions	19
IV.4.a.	Phase I - Summary statistics of key baseline measures and baseline equivalence across ePREP and waitlist control groups for couples completing relationship satisfaction measure (CSI-4) at 6-month follow-up	20
IV.4.b.	Phase I - Summary statistics of key baseline measures and baseline equivalence across OurRelationship and waitlist control groups for couples completing relationship satisfaction measures (CSI-4) at 6-month follow-up.....	21
IV.4.c.	Phase I - Summary statistics of key baseline measures and baseline equivalence across ePREP and OurRelationship groups for couples completing relationship satisfaction measure (CSI-4) at 6-month follow-up	22
V.1.	Covariates included in all Phase I impact analyses.....	27
V.2.a	Post-intervention estimated effects for ePREP vs. waitlist control using data from 6-month follow-up to address the primary research questions	27
V.2.b.	Post-intervention estimated effects for OurRelationship vs. waitlist control using data from 6-month follow-up to address the primary research questions	27
V.2.c.	Post-intervention estimated effects for ePREP vs. OurRelationship using data from 6-month follow-up to address the primary research questions	28
V.2.d.	Covariates included in Phase II impact analyses	29
V.2.e.	Post-intervention estimated effects for OurRelationship No Coach vs. Coach using data from 6-month follow-up to address the primary research questions.....	30
V.2.f.	Post-intervention estimated effects for OurRelationship Contingent Coach vs. Coach using data from 6-month follow-up to address the primary research questions.....	30

V.2.g. Post-intervention estimated effects for ePREP No Coach vs. Coach using data from 6-month follow-up to address the primary research questions	30
V.2.h. Post-intervention estimated effects for ePREP Contingent vs. Coach using data from 6-month follow-up to address the primary research questions	31
V.2.i. Post-intervention estimated effects for OurRelationship No Coach vs. Waitlist Control using data from 6-month follow-up to address the primary research questions	31
V.2.j. Post-intervention estimated effects for OurRelationship Contingent Coach vs. Waitlist Control using data from 6-month follow-up to address the primary research questions	32
V.2.k. Post-intervention estimated effects for ePREP No Coach vs. Waitlist Control using data from 6-month follow-up to address the primary research questions	32
V.2.l. Post-intervention estimated effects for ePREP Contingent Coach vs. Waitlist Control using data from 6-month follow-up to address the primary research questions	33
A.1. Data used to address implementation research questions	40
A.2. Key features of the impact analysis data collection	41
B.1. Phase II - Summary statistics of key baseline measures and baseline equivalence across study groups (OurRelationship No Coach and OurRelationship Coach), for couples completing relationship satisfaction measure (CSI-4) at 6-month follow-up	44
B.2. Phase II - Summary statistics of key baseline measures and baseline equivalence across study groups (OurRelationship Contingent Coach and OurRelationship Coach), for couples completing relationship satisfaction measure (CSI-4) at 6-month follow-up	45
B.3. Phase II - Summary statistics of key baseline measures and baseline equivalence across study groups (ePREP No Coach and ePREP Coach), for couples completing relationship satisfaction measure (CSI-4) at 6-month follow-up	46
B.4. Phase II - Summary statistics of key baseline measures and baseline equivalence across study groups (ePREP Contingent Coach and ePREP Coach), for couples completing relationship satisfaction (CSI-4) at 6-month follow-up	47
B.5. Phase II - Summary statistics of key baseline measures and baseline equivalence across study groups (OurRelationship No Coach and waitlist control), for couples completing relationship satisfaction measure (CSI-4) at 6-month follow-up	48

B.6.	Phase II - Summary statistics of key baseline measures and baseline equivalence across study groups (OurRelationship Contingent Coach and waitlist control), for couples completing relationship satisfaction measure (CSI-4) at 6-month follow-up	49
B.7.	Phase II - Summary statistics of key baseline measures and baseline equivalence across study groups (ePREP No Coach and waitlist control), for couples completing relationship satisfaction measure(CSI-4) at 6-month follow-up	50
B.8.	Phase II - Summary statistics of key baseline measures and baseline equivalence across study groups (ePREP Contingent Coach and waitlist control), for couples completing relationship satisfaction measure (CSI-4) at 6-month follow-up.....	51
E.1.	Description of ePREP Structure and Content	53
E.2.	Description of ePREP Structure and Content	54
E.3.	Detailed summary of implementation evaluation findings	57

Figures

III.1.	Phase 2 Design – a Limited SMART Design.....	13
A.1.	CONSORT diagram – Phase I (October 2016-December 2018).....	42
A.2.	CONSORT diagram – Phase II (January 2019-May 2020)	43

Impact Evaluation of OurRelationship and ePREP Online Couple Relationship Education Programs

I. Introduction

A. Introduction and study overview

Finding ways to effectively intervene with low-income couples is critically important. Low-income couples are more likely to experience adverse relationship outcomes, including low relationship satisfaction (Lundquist et al., 2014), intimate partner violence (Cunradi et al., 2002), and divorce (Bramlett & Moster, 2002). Unfortunately, low-income couples are also least likely to receive in-person interventions to improve their relationship or prevent breakup (Doss, 2014). Further, in-person relationship education programs are not always offered in places that are easily accessible to low-income couples, and even when available, there are multiple barriers to attendance, such as scheduling conflicts (especially around work or school), childcare difficulties, and transportation delays (Doss, 2014).

Delivering effective web-based relationship programs would help couples circumvent many of these common barriers to accessing relationship help, leading to improved relationship functioning among low-income couples. To test this possibility, the University of Miami study implemented a two-phase randomized controlled trial. In Phase I, 742 low-income couples were randomly assigned to either a Waitlist Control group, the OurRelationship program (Doss et al., 2016), or the ePREP program (Braithwaite & Fincham, 2011) and assessed for six months. In Phase II, 1,250 low-income couples were randomly assigned to receive the OurRelationship or ePREP interventions with varying levels of coach contact, or a Waitlist Control group. In both programs, couples assigned to OurRelationship or ePREP completed approximately 7 hours of online content.

OurRelationship and ePREP are two web-based programs that have different structure, content, and empirical basis, but show effectiveness in improving relationship outcomes among low-income couples (Doss et al., 2015; Braithwaite & Fincham, 2011, 2014). These web-based programs provide three important advantages over face-to-face Healthy Marriage and Relationship Education (HMRE) services: 1) They provide relationship services nationwide – even to rural areas and states; 2) They intervene effectively with distressed couples who have been underserved in previous relationship programs; and 3) They are cost-effective; maximizing the availability of services during and following the intervention.

The current impact evaluation provides the opportunity to 1) replicate previous findings around the effectiveness of web-based relationship education for improving relationship outcomes

among low-income couples; 2) compare two different web-based programs to one another; and 3) evaluate the utility of including different types of coach contact in these programs.

This report first describes the research questions of interest and services planned for participants including the program components, random assignment process, and data collection procedures for Phase I and II of the study. Next, the report reviews the actual implementation of the ePREP and OurRelationship programs, the impact of ePREP and OurRelationship on relationship satisfaction compared to a Waitlist Control group six months after baseline (Phase I), and the impact of ePREP and OurRelationship with varying levels of coach contact on relationship satisfaction compared to a Waitlist Control group six months after baseline (Phase II). Implications of these implementation and impact findings are discussed at the end of the report.

B. Primary research question(s)

This evaluation consists of a Phase I and Phase II. In brief, Phase I tested the relative effectiveness of the OurRelationship and ePREP online curricula, and Phase II tested the coaching component of the interventions in order to examine whether having coach contact from the beginning of the study, only after non-response, or not at all affected the impact of the programs on relationship satisfaction. When information in the current report differs by phase, we have provided separate sections for each phase. This study is registered with clinicaltrials.gov (Phase I: NCT02806635; Phase II: NCT03568565).

Phase I

1. Do couples assigned to OurRelationship have significantly higher relationship satisfaction 6 months after randomization compared to a Waitlist Control group?
2. Do couples assigned to ePREP have significantly higher relationship satisfaction 6 months after randomization compared to a Waitlist Control group?
3. Do couples assigned to OurRelationship and ePREP have significantly different relationship satisfaction 6 months after randomization compared to each other?

Phase II

1. Do couples assigned to OurRelationship No Coach have significantly lower relationship satisfaction 6 months after randomization compared to Coach OurRelationship?
2. Do couples assigned to OurRelationship Contingent Coach have significantly lower relationship satisfaction 6 months after randomization compared to Coach OurRelationship?
3. Do couples assigned to ePREP No Coach have significantly lower relationship satisfaction 6 months after randomization compared to Coach ePREP?
4. Do couples assigned to ePREP Contingent Coach have significantly lower relationship satisfaction 6 months after randomization compared to Coach ePREP?
5. Do couples assigned to OurRelationship No Coach have significantly higher relationship satisfaction 6 months after randomization compared to a Waitlist Control group?

6. Do couples assigned to OurRelationship Contingent Coach have significantly higher relationship satisfaction 6 months after randomization compared to a Waitlist Control group?
7. Do couples assigned to ePREP No Coach have significantly higher relationship satisfaction 6 months after randomization compared to a Waitlist Control group?
8. Do couples assigned to ePREP Contingent Coach have significantly higher relationship satisfaction 6 months after randomization compared to a Waitlist Control group?

C. Secondary research question(s)

There were no secondary research questions for this impact evaluation.

II. Intervention and counterfactual conditions

This section outlines the intended program and counterfactual conditions for the University of Miami impact evaluation, including the program content, planned dosage and implementation schedule, and the education and training of program staff. In Phase I, the program conditions included two web-based relationship education programs, OurRelationship and ePREP. Both programs included approximately 7 hours of online content and one hour of coach content – separated into four, 15-minute calls and spread throughout the duration of the program in order to reinforce program content (e.g., practicing skills learned) and troubleshoot participant issues (e.g., technological difficulties). The counterfactual condition was a Waitlist Control group wherein participants waited six months after enrollment before opting to complete the ePREP or OurRelationship program. Participants in the Waitlist Control group were asked to refrain from seeking relationship-focused services (either online or in-person) for the duration of the study.

In Phase II, the program conditions included OurRelationship and ePREP but with varying types of coach contact (No Coach, Coach, or Contingent Coach [i.e., coach contact only after non-response without a coach]). Non-response was indicated when couples did not complete required activities for 13 days or more. In this case, when deemed non-responsive, couples were re-randomized to receive a coach (Contingent Coach) or continue without a coach (No Coach). The counterfactual condition for Phase II was the same as Phase I (Waitlist Control).

A. Description of program as intended

1. Planned Dosage and Implementation Schedule

The planned dosage and implementation schedules for Phase I and Phase II were largely the same; however, given that the two phases included different samples of participants and randomization processes (see Random Assignment Process in section III.A for more details), the two phases will be discussed separately.

Phase I. Participants were low-income couples seeking relationship help who were randomly assigned to one of three conditions (ePREP, OurRelationship, or Waitlist Control). Both the ePREP and OurRelationship programs included online content (“Workshops”) delivered on a computer with a broadband connection or a smartphone and one hour of coach content – separated into four, 15-minute calls and spread throughout the duration of the program (our “Individualized Service Contacts”). These coach calls either occurred via a video conference feature such as Skype, FaceTime, or Google Hangouts or via a phone call. Both programs were designed to be completed over the course of 6-8 weeks.

OurRelationship. The OurRelationship program was adapted from Integrative Behavioral Couple Therapy, an effective in-person couple therapy (Christensen et al., 2004), and includes content on identifying a core relationship issue for the couple to work on, gaining an understanding of the problem (e.g., hidden stressors, communication patterns), and developing a plan for fixing the problem (e.g., acceptance and self-change, communication tips). Couples randomized to the OurRelationship program completed approximately 7 hours of online content

over 6 weeks and met with their coach after each phase of the program, as depicted in Table E.1 in the Appendix. Couples completed most activities individually by typing in responses to questions and prompts in the online program, then came together for structured conversations with their partner at the conclusion of each phase to discuss what they had written.

ePREP. The ePREP program is an online adaptation of the Prevention and Relationship Enhancement Program (PREP; Markman, Renick, Floyd, Stanley, & Clements, 1993) which provides couples with new information and skills to better their relationships, including healthy communication strategies, commitment, and fun and friendship. Couples randomized to the ePREP program completed 6 hours of online content, approximately 1-2 hours of additional homework over 6 weeks, and met with their coach every other week. The online content has a different theme each week (see Table H.2 in the Appendix) and couples are encouraged to watch the sections together. ePREP consists of psychoeducational presentations, videos of example couples, and questions for couples to discuss after the presentation.

More information on the content of specific activities for each program can be found in Tables E.1 and E.2 in the Appendix. Couples could end the online activities when they chose and pick up where they left off the next time they logged in to the program. As a result, session length varied. However, coaches indicated which content needed to be completed before the next scheduled call so that all couples had a completion goal to work towards. Given that these are the only two web-based interventions designed to strengthen couples' relationships that have been developed and tested, to our knowledge, the current impact report aims to compare and contrast each program's impact on relationship satisfaction compared to a Waitlist Control group six months after baseline.

Phase II. While the planned dosage and implementation schedule closely resembled that of Phase I, couples in Phase II were initially randomly assigned to one of five conditions (instead of three). All couples assigned to the Intervention Conditions (ePREP Coach, ePREP No Coach, OurRelationship Coach, or OurRelationship No Coach) were offered the ePREP and OurRelationship Workshops as described above. Couples assigned to ePREP Coach or OurRelationship Coach were offered the Individualized Service Contacts (four, 15-minute calls spread throughout their time in the program). Those in the No Coach conditions were not offered coaching initially, but if they did not engage with the workshops on their own (i.e., they didn't complete required activities for 13 days or more), they were re-randomized to either a "Contingent Coach" condition or not. Those who were re-assigned to coaching are then offered the same coaching as in the original coach groups. Participants randomly assigned (initially or upon re-randomization for non-responders) to No Coach did not have phone contact with a coach other than the initial call to obtain verbal informed consent. Instead, these participants completed the online activities with the aid of automated, contingent reminders and tailored feedback based on their responses. See Table II.1 for more information regarding the intervention and counterfactual conditions, as well as the target population for the study.

2. Education and Training of Staff

Education and training of staff was the same for both Phase I and Phase II (see Table II.2).

a. Hiring Requirements of Coaches

All coaches were required to have a bachelor's degree in Psychology or a related field. Additionally, all coaches were graduate students currently enrolled in a graduate program in a related field – usually a Ph.D. program in clinical psychology (at the University of Miami, Brigham Young University, or the University of Denver – our three partner universities). Additionally, some coaches were graduate students in the master's Marriage and Family Therapy program at the University of Miami or the Ph.D. program in clinical psychology at Florida International University or Nova Southeastern University (two universities in Miami). The program attempted to include diversity in the gender, race, and ethnicity of coaches.

b. Initial Training of Coaches

When joining the project, coaches completed both the ePREP and OurRelationship programs. Then, coaches participated in a two-day training where they learned about each of the types of contacts they have with couples, engaged in role playing to prepare for program coaching sessions (following the detailed scripts that all coaches use), and learned about general project procedures. Following this training, they observed previously trained coaches conducting calls with couples at different stages of the program. There were 7 different specific calls that a coach could have with a couple, depending on the program the couple was randomly assigned to: Welcome, OurRelationship Observe Phase, OurRelationship Understand Phase, OurRelationship Respond Phase, ePREP Phase 1, ePREP Phase 2 and ePREP Phase 3. Every new coach observed at least one of every type of call, then a lead intervention coach observed the new coaches during their first few contacts before allowing them to work with couples independently.

c. Ongoing Supervision of Coaches

Additionally, coaches had weekly group supervision meetings throughout the study where they were trained on different aspects of the OurRelationship or ePREP programs. Dr. Doss (developer of the OurRelationship program) or Dr. Braithwaite (ePREP partner; developer of the ePREP program) led these meetings. In these meetings, attendees listened to an audio recording of a coach call and the coaches and supervisors had an opportunity to comment. Recordings were chosen to represent common issues that arise and/or to highlight especially skillful coach delivery. Coaches also conducted quality assurance checks at these meetings, where they brought up special circumstances surrounding couples so that all coaches were familiar with how to provide high quality service in those situations and respond in the same way. Lastly, every week, coaches were required to submit an audio recording of one of their calls. The lead intervention developer for the specific program delivered on the call (ePREP or OurRelationship) reviewed these recordings in order to make sure that all coaches were giving the same quality of care and the correct level of care to participants. The intervention developers provided written feedback on all submitted recordings.

d. Target Population

In both phases, the target population was low-income couples who were married, engaged, or cohabiting with their partner for 6 months or more at the time of enrollment. In order to identify low-income couples specifically, couples were included if they have a household income within 200% of the federal poverty line. Additionally, couples were excluded if either participant reports intimate partner violence in the current relationship. Couples were recruited primarily via online advertising geared toward those who were seeking support for their relationships (e.g., counseling; see section III.A for more information on recruitment.)

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

Component	Curriculum and content	Dosage and schedule	Delivery	Target Population
Intervention				
Online Relationship Skills Workshops	OurRelationship or ePREP <u>online relationship education programs</u>	<u>ePREP</u> (6 hours online content + 1-2 hours homework), <u>OurRelationship</u> (7 hours online content), with flexible session lengths occurring for 6-8 weeks	Couples chose <u>the pace at which they completed the online content</u> . Couples could end the online activities when they chose and pick up where they left off the next time they logged in to the program.	Low-income couples (see III.A for further details on eligibility)
Coach Calls	Individualized service contacts to help couples: a) <u>practice skills learned from the online activities</u> and apply it to their own relationships; b) <u>complete the online content in a timely manner</u> ; and c) <u>assist with any technical difficulties</u> .	<u>Four, 15-minute calls</u> spread throughout the duration of the program	Phone call	<u>Phase I</u> : Low-income couples enrolled in ePREP or OR <u>Phase II</u> : Low-income couples enrolled in ePREP or OR Coach conditions or Contingent Coach conditions
Counterfactual (Waitlist Control Group)				
Waitlist Control	Couples instructed not to seek help for their relationship during the 6 months after enrollment; offered enrollment in their choice of the ePREP or OurRelationship programs after waiting period	Completed 4 surveys over the course of 6 month waiting period	NA	Low-income couples

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

Component	Education and initial training of staff	Ongoing training of staff
Intervention		
Individualized service contact ("Coach Calls")	Facilitators ("coaches") hold <u>at least a bachelor's degree</u> and were <u>currently enrolled in a graduate program</u> in a related field (as described in detail in text). Initial training consisted of: a) <u>completion of both online programs</u> (our "workshops"); b) <u>two days of initial training</u> (where they reviewed call scripts, listened to recordings of calls, and role-played those calls); c) <u>listening live to a lead coach conduct a call</u> ; d) <u>obtaining live supervision</u> from a lead coach on each type of call (for both curricula).	Facilitators participate in <u>required weekly supervision</u> (1 hour), where an audio recording of a call is reviewed and discussed. Coaches also have an opportunity to receive <u>group supervision</u> on specific couples on their caseload that are experiencing difficulties. In addition, each coach <u>submits a recording of one call a week to the lead coach</u> (usually a call that was especially challenging). The lead coach listens to the recording and provides written feedback.
Counterfactual (Waitlist Control Group)		
No services provided during waiting period	Not applicable	Not applicable

B. Description of counterfactual condition as intended

1. Phase I Services for Waitlist Control Group

All couples attended an initial, 15-minute call with a coach; the same coaches conducted the calls for the intervention and Waitlist Control groups. On this call, the project was further explained to couples (they already provided informed consent at this point) and couples had the opportunity to ask any questions. After providing verbal consent, couples were then informed of their random assignment. Couples randomized to the Waitlist Control group were reminded of the requirements of this condition (including not seeking other services for their relationship during this period).

Project staff contacted couples in the Waitlist Control group via email (and then by phone, if initially non-responsive) by one, two, four, and six months following randomization and provided them a link to an online survey. Couples did not have any other contact with their coaches or project staff during this period. At the end of six months, couples were contacted and provided access to their choice of online program (ePREP or OurRelationship) if they were still interested in doing so.

2. Phase II Services for Waitlist Control Group

All Waitlist Control couples received the same contacts with project staff and coaches as in Phase I. Phase II couples in the Waitlist Control group were asked to refrain from seeking relationship-focused services (either online or in-person) for the duration of the study.

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

Table II.3 outlines research questions regarding the implementation of the ePREP and OurRelationship programs, including fidelity to the program, program dosage received by participants, quality of staff-participant interactions, couples' feedback of the program, and outside services utilized by couples during the intervention period. These findings will provide context and aid interpretation of the impact analysis findings for couples assigned to the ePREP or OurRelationship programs.

Table II.3. Research questions about implementation of OurRelationship/ePREP and the Waitlist Control Group

Implementation element	Research question
Intervention Group Questions	
Fidelity	<ul style="list-style-type: none"> • Were all intended intervention components offered and for the expected duration? • Were there adaptations to the workshops over the course of the study? • Were individual service contacts delivered reliably across providers to those in the intervention group?
Dosage	<ul style="list-style-type: none"> • What percentage of couples who enrolled in the intervention completed 100% of the intervention (i.e., workshops, individual service contacts)? • What percentage of couples who enrolled in the intervention completed 2/3 of the intervention (i.e., workshops, individual service contacts)?
Quality	<ul style="list-style-type: none"> • What was the quality of staff—participant interactions during the intervention (i.e., workshops and individual service contacts)?
Engagement	<ul style="list-style-type: none"> • What was intervention couples' overall feedback on the program?
Context	<ul style="list-style-type: none"> • What outside services were utilized by couples in the intervention group during the evaluation period?
Counterfactual (Waitlist Control Group) Questions	
Fidelity	<ul style="list-style-type: none"> • Was the initial Welcome call offered to all couples in the control group and for the expected duration? • What outside services were utilized by couples in the control group during the evaluation period?

III. Study design

This section provides an overview of the sample formation, including eligibility criteria, consent process, and random assignment process for both Phase I and II.

A. Sample formation and research design

1. Sample formation

In both Phase I and Phase II, participants from across the United States were directed to the OurRelationship.com website primarily through a combination of organic search results and paid advertising on Google. For example, two of the most common search terms that led couples to the site are “online marriage help” and “free marriage counseling”. Other recruitment methods included social media (primarily Facebook) and word of mouth from previous participants. On the OurRelationship website, couples read about the nature of the online programs, the accompanying coach calls, and the requirements of the local evaluation. If couples had questions, they were able to contact project staff using a contact form on the website, by email, or by calling the project office.

2. Sample eligibility criteria

In both Phase I and Phase II, interested individuals completed an online informed consent form and a subsequent eligibility survey. Couples were eligible if they reported on an online survey that they were married, engaged, or cohabiting with their partner for 6 months or more; had a household income within 200% of the federal poverty line; lived in the United States; were between the ages of 18 and 64 (inclusive); were able to fluently read and write in English; had high speed or 3G internet access; agreed to forego other couple treatment for the next 6 months; and had not previously participated in PREP, ePREP or the OurRelationship programs. Additionally, couples were excluded if either participant answered affirmatively to any of the following behaviors used to screen for intimate partner violence (IPV; in the past 6 months) from their partner: 1) Choked me 2) Repeatedly punched me during a fight or beat me up 3) Threatened me with (or used) a gun or knife, or 4) Physically forced me to have sex when I didn't want to. Furthermore, couples were excluded if either partner reported in the past 6 months that: 5) he/she was “quite afraid”, “very afraid”, or “extremely afraid” that their partners would physically hurt them during an argument. If the first member of the couple is eligible, they are provided instructions to send the screening survey to their partner. The second partner completed an identical screening survey, with the exception of the income questions given that we already had the total household income from the first partner. Ineligible participants were offered appropriate referrals (e.g., other online programs, self-help books or, if violence was endorsed, resources such as the National Domestic Violence Hotline).

3. Consent process

The study was approved by the University of Miami IRB on 6/27/2016 (Protocol 2016-0451; “Web-based Marriage Preparation/Enrichment and Divorce Prevention for Low-Income

couples”). In addition to minor modifications, the study’s approval was renewed on an annual basis on 5/22/2017, 3/26/2018, 2/7/2019, and 1/6/2020.

For both Phase I and Phase II, consent was obtained before any measures are completed. When couples visited the study’s website, they read about the ePREP and OurRelationship programs and what participation in the study entailed. If a couple decided that they wanted to participate in the web-based programs, they clicked a hyperlink that took them to the online screening survey (“Applicant Characteristics”). The first page of the screening survey was an online informed consent form. After reading the informed consent form, the participant checked a box agreeing to participate. The participant was not able to progress to the rest of the survey until that box was checked.

After both members of the couple provided online consent and completed both baseline surveys, they were scheduled for a 15-minute phone call or videoconference with a Coach. On this call, the Coach briefly reviewed the components of the study and answered any questions the couple had. Couples were then given a chance to ask any questions they had about the research requirements. Then, they were asked to verbally agree to key requirements of the study.

Consent procedures were identical for couples in the intervention groups and the Waitlist Control group. Couples were only told of their randomized condition after they had both provided verbal consent (and thus are formally enrolled in the study).

1. Random Assignment Process

2. Phase I. Couples were randomly assigned into one of three groups (ePREP, OurRelationship, or a Waitlist Control group) using the following process:

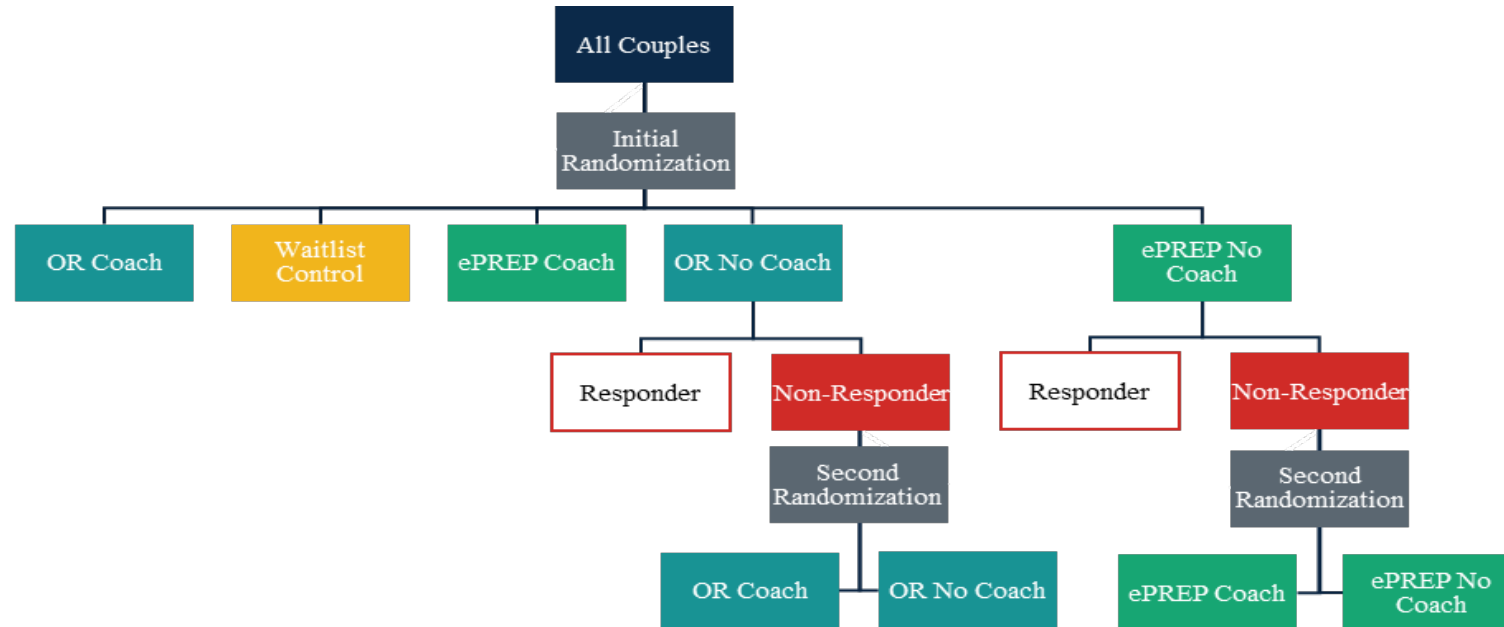
1. After both members of a couple provided online consent (see Consent Process above), completed the baseline measures, and confirmed their initial appointment with their coach, the Participant Coordinator (PC) examined each partner’s scores on a standard measure of relationship satisfaction (from the baseline survey). The PC then stratified the couple into either a “distressed” group (at least one partner in the clinically distressed range of relationship satisfaction on the 4-item version of the Couple Satisfaction Inventory) or “non-distressed” group (neither partner in the clinically distressed range on relationship satisfaction).
2. Within the two stratification groups, the PC then selected the next available randomized assignment from a pre-generated, randomized list in Excel (described in more detail below). Within each of the two stratification categories, couples had a 1/3 chance of being assigned to any of the three conditions. The random assignment was registered in the central project tracking log and in the couple’s individual online record (which was used in the next step by the coach).
3. Once both members of a couple provided verbal consent on their initial call with the PC (see Consent Process above), the coach looked up the couple’s random assignment in their online record (which was entered by the PC in the previous step), informed the couple about their

assignment, and reminded the couple about the expectations for that condition. This initial call followed a strict script both before and following randomization.

4. If one or both members of a couple did not provide verbal consent (or did not attend an initial call), the coach notified the PC and that couple was removed from the randomized assignment list. That slot then became available for the next couple who was eligible to be randomized.

The two random assignment lists (one for each of the two stratification categories) used by the PC were generated in Excel using a random number generator. Specifically: a) 1,000 random numbers were generated using the RAND() function; b) in a separate column, using the RANK function, each group of three sequential random numbers (e.g., rows 4, 5, and 6) was given a rank relative to the other two numbers (yielding non-repeating integers of 1, 2, and 3 in a random order). Each of these three numbers was then assigned to a randomized condition (e.g., 3 = control). The randomization process yielded three study groups with similar sample sizes ($n_{\text{ePREP}} = 247$ couples, $n_{\text{OR}} = 248$ couples, $n_{\text{waitlist}} = 247$ couples).

Phase II. As noted above, Phase I and Phase II were separate studies with different participant samples and randomization processes that were independent from one another. Couples in Phase II were randomly assigned into one of five groups (ePREP Coach, OurRelationship Coach, ePREP No Coach, OurRelationship No Coach, or a Waitlist Control group) using the same methods as in Phase I; however, the randomization for Phase II was completed separately from Phase I. Eighteen percent of couples were assigned to the Waitlist Control group, 18% to the ePREP Coach group, 18% to the OurRelationship Coach group, and 23% to the ePREP No Coach group, and 23% to the OurRelationship No Coach group. Phase II also utilized a limited SMART (Sequential, Multiple-Assignment, Randomized Trial; Lei et al., 2012) design wherein couples initially randomized to complete an intervention without a coach but who were deemed as “non-responsive” to the program were re-randomized (with 50% re-randomized to continue without a coach and 50% re-randomized to receive a coach). Specifically, couples were defined as “non-responsive” if either member did not complete a block of online activities comprising approximately 1/6th of the online program within 13 days. Couples could be classified as non-responsive during any of the six program blocks. See Figure 1 for a depiction of the Phase 2 design. Our initial randomization plan assumed that 50% of couples initially assigned to the “No Coach” condition would be categorized as a “non-responder” and would have yielded approximately 17.5% of couples per group in each of the planned analyses. However, approximately 70% of these couples were actually categorized as a “non-responsive,” yielding somewhat unbalanced groups for this comparison.

Figure III.1. Phase 2 Design – a Limited SMART Design

SMART designs include multiple, separate stages wherein participants were randomized to different treatment options (in this case the first stage involved initial randomization, and the second stage involved re-randomization of non-responsive participants). Randomly assigning participants to Coach/No Coach followed by re-randomizing non-responders in the No Coach category to receive a coach is necessary to demonstrate the causal effect of coach contact on program completion rates and effects. If we did not randomize and instead simply looked at the correlation between amount of coach contact and program outcomes/completion rates, the correlation would likely be minimal (and perhaps even negative) because some highly-motivated couples would quickly and effectively complete the program (requiring minimal coach contact) while less motivated couples (who would be less likely to complete or benefit from the program) would require more coach contact. Finally, randomizing non-responders in the No Coach condition has two benefits: a) it allows us to test whether adding a coach for those who are non-responsive is a useful approach and allow for more nuanced discussion of the costs and benefits of including coaching in relationship education; b) if adding a coach is indeed helpful, it will allow more couples to benefit from our services. SMART designs have been utilized in a number of federally-funded intervention trials and appropriate statistical analysis and power characteristics are well-defined. See <https://methodology.psu.edu/ra/adap-inter/research> for an overview.

B. Data collection

This section outlines the University of Miami study's data collection procedures; first discussing the implementation evaluation and then the impact evaluation.

1. Implementation analysis

In both Phase I and II, all survey items were collected via internal tracking logs, observations, the federally required Information, Family Outcomes, Reporting, and Management survey (nFORM), and OurRelationship and ePREP server logs. Specifically, the OurRelationship and ePREP website hosting companies monitored the website servers in order to ensure that all workshop components were available to participants at all times. In addition, internal tracking logs included encrypted tracking systems of individual service contacts with participants by project staff, and what percentage of couples enroll in the intervention. See Table A.1 in the Appendix for more information regarding data collection for the implementation analysis.

2. Impact analysis

In both Phase I and II, all survey items were collected via Qualtrics, including the required ACF nFORM survey items. In other sites, nFORM surveys were typically administered in real-time on a tablet. Given that participants completed the program online, program staff inserted the required nFORM survey items into the Qualtrics surveys that corresponded to the appropriate nFORM time point (e.g., Applicant Characteristics and Pre-Test nFORM items were included in the Baseline Qualtrics survey and were not altered in any way from their original ACF formatting). Local evaluator staff then manually entered participant nFORM item responses from Qualtrics into nFORM so that these records were accessible to ACF. Local evaluator staff met weekly to review data entry procedures, conduct quality checks, and resolve any data entry issues. Only the survey responses collected in Qualtrics were used in the implementation and impact analyses, not the manually entered data in nFORM.

In the Intervention and Waitlist Control groups, participants were administered five surveys: Baseline, Mid (1 month after Baseline), Post (2 months after Baseline), 4-Month Follow-Up (4 months after Baseline), and 6-Month Follow-Up (6 months after Baseline). Demographic information was collected only at Baseline. All surveys included items on relationship functioning, parenting functioning, and individual functioning. Participants were able to complete each survey until the point where they were closer in time to the next scheduled survey time point than they were to the original survey time point. For example, after three months following randomization, couples were closer in time to the four-month follow-up survey than they were to the expected completion date of the post-intervention survey (approximately 2 months following randomization); therefore, at three months, program staff stopped contacting participants for post-intervention data. See Table A.2 in the Appendix for more information on data collection for the impact analysis. Note that in order to reduce the scope for the current report, only Baseline and 6-Month Follow-Up data were included.

IV. Analysis methods

A. Analytic sample

The analytic sample presented in this report is based only on the subsample of couples for which both partners completed the CSI-4 at baseline and the 6-month follow-up. The final sample for Phase I includes 496 individuals who participated in OurRelationship, 494 individual who participated in ePREP, and 494 individuals in the Waitlist Control group (see Table IV.1a). Based on the What Works Clearinghouse conservative attrition standards, the overall and differential attrition for each outcome measure suggests that Phase I is a low-attrition RCT.

The final sample for Phase II includes 410 individuals assigned to OurRelationship Coach, 262 individuals assigned to OurRelationship Contingent Coach, 264 individuals assigned to OurRelationship No Coach, 397 individuals assigned to ePREP Coach, 259 individuals assigned to ePREP Contingent Coach, 259 individuals assigned to ePREP No Coach, and 405 individuals assigned to the Waitlist Control group (see Table IV.1b and IV.1c).¹ While the overall attrition for Phase II is low, the differential attrition between some groups suggests high attrition according to the What Works Clearinghouse conservative attrition standards. See Figures A.1 and A.2 in the Appendix for Phase I and II CONSORT diagrams.

¹ At the writing of this report, 6-month follow-up data collection for Phase II was not yet complete; therefore, analyses were based on participants who completed their 6-month follow-up by May 1, 2020 (N=1,128; 90% of the sample), rather than the end of the study (N=1,250).

Table IV.1a. Phase I - Individual sample sizes for the analytic sample by intervention status

Number of couples	Our-Relationship sample size	ePREP sample size	Waitlist Control sample size	Total sample size	Total response rate	Our-Relationship response rate	ePREP response rate	Waitlist Control response rate
Assigned to condition	496	494	494	1,484	n.a.	n.a.	n.a.	n.a.
Contributed a baseline survey	496	494	494	1,484	100%	100%	100%	100%
Contributed to 6-month follow-up survey (6 months after baseline)	411	409	471	1,291	87%	83%	83%	95%
Both partners contributed to BL and 6-month follow-up survey (CSI-4)*	346	328	412	1,086	73%	70%	66%	83%

n.a. = not applicable.

*indicates cases in which both partners completed CSI-4 at baseline and 6-month follow-up and accounts for item non-response and any other analysis restrictions.

Table IV.1b. Phase II - Individual sample sizes for the analytic sample by intervention status – OurRelationship

Number of couples	OR Coach sample size	OR Contingent Coach sample size	OR No Coach sample size	Waitlist Control sample size	Total sample size	Total response rate	OR Coach response rate	OR Contingent Coach response rate	OR No Coach response rate	Waitlist Control response rate
Assigned to condition	410	262	264	405	1,341	n.a.	n.a.	n.a.	n.a.	n.a.
Contributed a baseline survey	410	262	264	405	1,341	100%	100%	100%	100%	100%
Contributed to 6-month follow-up survey (6 months after baseline)	302	200	187	356	1,045	78%	74%	76%	71%	88%
Both partners contributed to BL and 6-month follow-up survey (CSI-4)*	262	160	160	324	930	69%	64%	61%	61%	80%

Notes: Sample sizes and response rates are based only on those eligible to take the 6-month follow-up survey as of May 1, 2020. Total individual sample size across OurRelationship and ePREP was 2,256 but tables are presented separately for each program given space restrictions.

*indicates cases where both partners completed CSI-4 at baseline and 6-month follow-up and accounts for item non-response and any other analysis restrictions

Table IV.1c. Phase II - Individual sample sizes for the analytic sample by intervention status – ePREP

Number of couples	ePREP Coach sample size	ePREP Contingent Coach sample size	ePREP No Coach sample size	Waitlist Control sample size	Total sample size	Total response rate	ePREP Coach response rate	ePREP Contingent Coach response rate	ePREP No Coach response rate	Waitlist Control response rate
Assigned to condition	397	259	259	405	1,320	n.a.	n.a.	n.a.	n.a.	n.a.
Contributed a baseline survey	397	259	259	405	1,320	100%	100%	100%	100%	100%
Contributed to 6-month follow-up survey (6 months after baseline)	296	184	210	356	1,046	79%	75%	71%	81%	88%
Both partners contributed to BL and 6-month follow-up survey (CSI-4)*	246	158	184	338	926	70%	62%	61%	71%	83%

Notes: Sample sizes and response rates are based only on those eligible to take the 6-month follow-up survey as of May 1, 2020. Total individual sample size across OurRelationship and ePREP was 2,256 but tables are presented separately for each program given space restrictions.

*indicates cases where both partners completed CSI-4 at baseline and 6-month follow-up and accounts for item non-response and any other analysis restrictions

B. Outcome measures

The primary impact analyses for both Phase I and Phase II evaluated differences in relationship satisfaction by the intervention groups and waitlist control group. To measure relationship satisfaction, the sum of the Couple Satisfaction Index (CSI-4) was calculated at baseline and six months after baseline, with higher scores indicating greater relationship satisfaction (see Table IV.2).

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

Outcome measure	Description of the outcome measure	Source	Timing of measure
Relationship Satisfaction	Participants were asked to rate how happy they were in their relationship using the four-item version of the Couple Satisfaction Index. One item was rated on a 0-6 scale (0=Extremely Unhappy, 6=Perfect) and three items were rated on a 0-5 scale (0=Not at all true, 5=Completely), with higher scores representing greater relationship satisfaction. An example item is: "I have a warm and comfortable relationship with my partner." A participant score was created by summing the four items, Phase I alpha = .92; Phase II alpha=.93.	The four-item version of the Couple Satisfaction Index (CSI-4; Funk & Rogge, 2007)	Six months after baseline

C. Baseline equivalence and sample characteristics

1. Phase I

For Phase I, the effect size difference was calculated between the intervention groups and Waitlist Control groups at baseline on relationship satisfaction, as well as key characteristics such as gender, race/ethnicity, and marital status, and these variables were controlled for as necessary. Although Phase I is considered a low-attrition RCT, variables that did not demonstrate baseline equivalence were controlled for in order to improve efficiency of the impact estimates and further minimize any risk of bias. Hedges' g was used to calculate the effect sizes for continuous variables by dividing the differences in means by the weighted pooled standard deviation. Effect sizes for dichotomous variables are Cox indices in comparable units.

At baseline in Phase I, there were effect size differences between .05 and .25 in key characteristics when comparing ePREP to Waitlist Control Group, OurRelationship to Waitlist Control Group, and ePREP to OurRelationship (see Tables IV.4.a-c); thus, the Phase I impact analyses controlled for race, ethnicity, and marital status.

Table IV.4.a. Phase I - Summary statistics of key baseline measures and baseline equivalence across ePREP and waitlist control groups for couples completing relationship satisfaction measure (CSI-4) at 6-month follow-up

Baseline measure	ePREP <i>M (SD)</i> or %	Waitlist Control <i>M (SD)</i> or %	ePREP vs. Waitlist Control regression coefficient (p-value)	Effect size
Female (%)	52	52	b = 0.017 (p = .910)	0.00
Race/ethnicity (%)				
Hispanic/Latino	10	12	b = -0.223 (p = .729)	0.12
Caucasian/White	66	60	b = 0.591 (p = .145)	0.16
Marital Status (%)	59	59	b = 0.01 (p = .993)	0.00
Relationship satisfaction (range: 0 to 21)	9.59 (4.62)	9.53 (4.77)	b = 0.057 (p = .893)	0.01
Sample size	328	412	n.a.	n.a.

n.a. = not applicable.

Notes: p-values are included in parentheses. Effect sizes for continuous variables are Hedges' g, calculated by dividing the differences in means by the weighted pooled standard deviation. Effect sizes for dichotomous variables are Cox indices in comparable units.

Table IV.4.b. Phase I - Summary statistics of key baseline measures and baseline equivalence across OurRelationship and waitlist control groups for couples completing relationship satisfaction measures (CSI-4) at 6-month follow-up

Baseline measure	OurRelationship <i>M</i> (<i>SD</i>) or %	Waitlist Control <i>M</i> (<i>SD</i>) or %	OurRelationship versus Waitlist Control regression coefficient (p-value)	Effect size
Female (%)	53%	52%	b = 0.026 (p = .856)	0.02
Race/ethnicity (%)				
Hispanic/Latino	11%	12%	b = -0.136 (p = .828)	0.06
Caucasian/White	69%	60%	b = 0.926 (p = .023)	0.24
Married (%)	54%	59%	b = 0.228 (p = .804)	0.12
Relationship satisfaction (range: 0 to 21)	9.58 (4.52)	9.53 (4.77)	b = 0.041 (p = .922)	0.01
Sample size	346	412	n.a.	n.a.

n.a. = not applicable.

Notes: p-values are included in parentheses. Effect sizes for continuous variables are Hedges' g, calculated by dividing the differences in means by the weighted pooled standard deviation. Effect sizes for dichotomous variables are Cox indices in comparable units.

Table IV.4.c. Phase I - Summary statistics of key baseline measures and baseline equivalence across ePREP and OurRelationship groups for couples completing relationship satisfaction measure (CSI-4) at 6-month follow-up

Baseline measure	ePREP Mean (SD) or %	OurRelationship mean Mean (SD) or %	ePREP versus OurRelationship regression coefficient (p-value)	Effect size
Female (%)	52%	53%	b = 0.043 (p = .780)	0.02
Race/ethnicity (%)				
Hispanic/Latino	10%	11%	b = -0.084 (p = .908)	0.06
Caucasian/White	66%	69%	b = -0.427 (p = .427)	0.08
Married (%)	59%	54%	b = 0.241 (p = .810)	0.12
Relationship satisfaction (range: 0 to 21)	9.59 (4.62)	9.58 (4.52)	b = 0.016 (p = .970)	0.00
Sample size	328	346	n.a.	n.a.

n.a. = not applicable.

Notes: p-values are included in parentheses. Effect sizes for continuous variables are Hedges' g, calculated by dividing the differences in means by the weighted pooled standard deviation. Effect sizes for dichotomous variables are Cox indices in comparable units.

2. Phase II

For Phase II, the effect size difference was calculated between the intervention groups and Waitlist Control groups at baseline on relationship satisfaction, as well as key characteristics such as gender, race/ethnicity, and marital status, and these variables were controlled for as necessary. Certain comparisons in Phase II were considered to have high attrition due to differential attrition, including ePREP Coach versus ePREP No Coach, ePREP Coach versus Waitlist Control, ePREP Contingent Coach versus Waitlist Control, OurRelationship No Coach versus Waitlist Control, and OurRelationship Contingent Coach versus Waitlist Control (see Tables V.1b-c). Therefore, variables that failed to demonstrate baseline equivalence were controlled for in order to statistically adjust for the inequivalence. Hedges' g was used to calculate the effect sizes for continuous variables by dividing the differences in means by the weighted pooled standard deviation. Effect sizes for dichotomous variables are Cox indices in comparable units.

At baseline in Phase II, there were effect size differences between .05 and .25 in key characteristics across comparison groups (e.g., OurRelationship No Coach vs. Coach; see Tables B.4-8 in the Appendix); thus, Phase II impact analyses controlled for race, ethnicity, marital status, gender, and/or baseline levels of relationship satisfaction. In addition, in comparisons of baseline levels of marital status between OurRelationship No Coach versus Coach and OurRelationship No Coach versus Waitlist Control, the effect size difference between groups

was greater than .25 (see Tables B.4, & B.8 in the Appendix). Given that the comparison of OurRelationship No Coach and Waitlist Control exhibited high differential attrition and differences in baseline marital status outside of the normal range, we used a propensity score matching procedure to select equivalent samples in each group that were maximally comparable. This method ensures that the only meaningful difference between the treatment and control group is the treatment status and not baseline levels of the covariates (see Section B in the Appendix). Given that the comparison of OurRelationship No Coach versus OurRelationship Coach exhibited low differential attrition, marital status was controlled for in this impact analysis, but propensity score weighting was not used. See Tables B.1-B.8 in the Appendix for summary statistics of key baseline measures and baseline equivalence across study groups in Phase II.

V. Findings and Estimation Approach

A. Implementation evaluation

Key Findings:

Results demonstrate that the ePREP and OurRelationship programs were implemented with fidelity and participants received a reasonably high dosage of program content in both Phase I and Phase II. In Phase II, the findings indicate that being assigned a coach may be associated with overall higher completion rates. Staff-participant interaction quality was only measured during Phase I and indicated high quality staff-participant interactions during the individual service contacts for both ePREP and OurRelationship. In addition, there was overall high participant engagement in both programs across Phase I and II with the minority of participants seeking outside relationship resources while enrolled in either program. Regarding fidelity for the Waitlist Control group, who was asked to refrain from seeking relationship-focused services for the duration of the study, a small portion of participants sought outside services during Phase I. Fidelity for the Waitlist Control group was not measured in Phase II.

Fidelity for intervention group. For both Phase I and Phase II, all intervention programs were delivered with fidelity over the course of the service, and participants had access to the OurRelationship and ePREP servers for workshop participation 99.99% of the time. Adaptations to the program occurred during Phase I in order to increase participant engagement in Phase II. First, a Spanish version of the program was released in order to accommodate Spanish-speaking couples in Phase II. However, only a small portion of couples in Phase II enrolled in the Spanish program ($n=16$). In addition, couples without a coach in Phase II received a complex series of emails contingent on completion or non-completion of activities as well as responses to the program. For example, when a couple indicated that a program conversation with their partner had not gone well, they received an automated email with suggestions on how to recover from the argument and re-engage with the program. There was only one deviation in service delivery by coaches, which occurred during Phase I (see Table E.3), suggesting that coaches engaged in the service delivery of Phase I and Phase II with fidelity.

Dosage. Although program completion rates did not reach the study target (75%), they were reasonably high for both intervention programs in Phase I, and the completion rates were similar across the two programs (OurRelationship (OR) = 68%; ePREP = 69%; $\chi^2(1) = .168, p = .682$). For ePREP, the overall completion rate in Phase II was similar to Phase I ($\chi^2(1) = .007, p = .935$), and the completion rates across the levels three levels of coach contact did not significantly differ, though there was a trend for the No Coach condition to be lower than the others (ePREP Coach = 69%, ePREP Contingent Coach = 72%, ePREP No Coach = 63%; $\chi^2(2) = 5.48, p = .064$). For OurRelationship, the completion rate in Phase II (OR Coach = 63%) was somewhat lower than Phase I but did not reach statistical significance ($\chi^2(1) = 2.83, p = .092$). As demonstrated in comparisons of OurRelationship Coach to OurRelationship Contingent Coach and No Coach, not offering a coach notably reduced the program completion rate (OR Contingent Coach = 53%; OR No Coach = 53%; $\chi^2(2) = 11.01, p = .004$). Taken together, these

findings suggest that participants received a moderate dosage of program content in Phases I and II and that being assigned a coach may be associated with overall higher completion rates.

Quality. In Phase I, staff-participant interactions during the individual service contacts were of high quality for both programs as suggested by the average alliance rating, calculated as the average of three alliance questions completed by participants after each coach call. On a 1 (not at all true) to 5 (extremely true) scale, couples in both programs on average reported that the individual service contacts were helpful, made them felt heard, understood and respected, and that what they wanted to talk about what was covered during the service ($M_{OR} = 3.94$, $SD = .93$; $M_{ePREP} = 4.18$, $SD = .86$). Across the two programs, using couple level data (averaged across partners), the staff-participant interaction quality was significantly higher in ePREP than OR ($t(245) = 2.04$, $p = .042$); however, participants in both programs reported overall high-quality staff-participant interactions. Data on staff-participant interaction quality was not gathered for Phase II.

Engagement. In Phases I and II, intervention participants reported a 1 (a lot), 2 (some), or 3 (not at all) in response to how much the program helped them. Mean scores suggest participants viewed both the ePREP and OurRelationship programs as helpful across Phase I and Phase II (see Table E.3 for means and standard deviations for each group within Phase I and Phase II). These ratings did not differ across different intervention conditions in either Phase I ($t(442) = -1.40$, $p = .162$) or Phase II ($F(5,918) = .28$, $p = .923$). Moreover, intervention couples rated on a 1 (no, definitely not) to 4 (yes, definitely) scale indicating that they would recommend the program to a friend in need of similar help. Mean scores suggest that participants would recommend both the ePREP and OurRelationship programs to a friend (see Table E.3 for means and standard deviations for each group within Phase I and Phase II). There were no significant differences across intervention conditions in either Phase I ($t(442) = -.07$, $p = .943$) or Phase II ($F(5,918) = 1.64$, $p = .146$). Taken together, these findings suggest overall high engagement with ePREP and OurRelationship across study phases, regardless of coach contact.

Context for intervention group. In Phase I, during the intervention period, 21.8% of the OR couples and 21.7% of the ePREP couples reported having used at least one outside resource. The most highly endorsed resources included self-help books or DVDs (10.7% OR, 8.7% ePREP), couples therapy (6.7% OR, 7.3% ePREP), and individual therapy (4.4% OR, 4.7% ePREP). In Phase II, during the intervention period, across all intervention conditions, less than 2% of couples reported having gone to couples therapy or any type of relationship education program other than the OurRelationship or ePREP online programs in the past two months.

Fidelity for control/comparison group. All Waitlist Control couples (100%) in both Phase I and Phase II attended the single required individual service contact. In Phase I, 13.4%, 14.6% and 16.4% of the couples reported having used outside services at 2 (post), 4, and 6 months after randomization, respectively.

For both Phase I and Phase II, percentages and averages were generated based on quantitative data such as the amount of time the program was made available online, the proportion of program completed, and participants' evaluation of the alliance and the program. Percentages

and averages were compared across intervention groups using independent sample t-test, one-way ANOVA and contingency table chi-square test when appropriate to determine whether the dosage and implementation quality differed across intervention condition (see Table E.3 in the Appendix for a detailed description of Implementation Findings).

1, Key limitations of the implementation data

1. In both Phase 1 and Phase 2, definitions of activity completion across ePREP and OR were not equivalent due to program requirements. Specifically, in ePREP, only one member of the couple need to complete the required activities to be marked as completing the program whereas in OR, both members of the couple need to complete the required activities to be marked as complete. This is likely part of the explanation for why ePREP completion rates were higher than OR. However, we do not have the data to test this hypothesis.
2. In Phase 2, data on outside services utilized by couples in the control group were not collected during the evaluation period. As a result, if gains are observed in waitlist couples during the evaluation period, we will not be able to tell if any of those may be due to outside services.
3. In Phase 2, data on the quality of staff – participant interactions during the individual service contacts was not collected. Thus, we are not able to examine whether the quality of individual service contact differed across different intervention conditions and whether it explained any of the differences observed in completion rates across intervention conditions.

B. Primary impact evaluation

All impact analyses for Phase I and Phase II were conducted using an Intent to Treat (ITT) framework, which is the most rigorous evaluation approach for RCTs. That is, participants' responses were analyzed based on their original group assignment, regardless of actual services received, to preserve the validity of the random assignment.

Key Findings – Phase I:

Phase I results suggest that both the OurRelationship and ePREP groups had significantly higher relationship satisfaction six months after randomization compared to a Waitlist Control group and after controlling for race, ethnicity, and marital status. In addition, the OurRelationship and ePREP groups did not significantly differ from one another in relationship satisfaction six months after randomization.

Table V.1. Covariates included in all Phase I impact analyses

Covariate	Description of the covariate
Hispanic/Latino	Ethnicity is Hispanic or Latino (1) versus not Hispanic or Latino (0)
Caucasian/White	Race is Caucasian/White (1) versus not Caucasian/White (0)
Marital Status	Couple was married (1) versus unmarried (0) at baseline

Table V.2.a Post-intervention estimated effects for ePREP vs. waitlist control using data from 6-month follow-up to address the primary research questions

Outcome measure	ePREP Mean (SD) or %	Waitlist Control Mean (SD) or %	ePREP versus Waitlist Control regression coefficient (p-value of difference)	Effect size
Relationship satisfaction	13.35 (9.22)	10.93 (8.24)	b = 2.42 (p < .001)	0.28
Sample Size	328	412	-	-

Source: Six-month follow-up surveys administered six months after randomization

Notes: Analyses are linear mixed effects models with individuals nested within couple and include covariates. Effect size is Hedges' g, calculated by dividing the differences in estimated marginal means by the weighted pooled standard deviation (converted from model standard error).

Table V.2.b. Post-intervention estimated effects for OurRelationship vs. waitlist control using data from 6-month follow-up to address the primary research questions

Outcome measure	OurRelationship Mean (SD) or %	Waitlist Control Mean (SD) or %	OurRelationship versus Waitlist Control regression coefficient (p-value of difference)	Effect size
Relationship Satisfaction	13.26 (9.19)	10.95 (8.41)	b = 2.32 (p < .001)	0.26
Sample Size	346	412		

Source: Six-month follow-up surveys administered six months after randomization

Notes: Analyses are linear mixed effects models with individuals nested within couple and include covariates. Effect size is Hedges' g, calculated by dividing the differences in estimated marginal means by the weighted pooled standard deviation (converted from model standard error).

Table V.2.c. Post-intervention estimated effects for ePREP vs. OurRelationship using data from 6-month follow-up to address the primary research questions

Outcome measure	ePREP Mean (SD) or %	OurRelationship Mean (SD) or %	ePREP versus OurRelationship regression coefficient (p-value of difference)	Effect size
Relationship Satisfaction	13.36 (8.88)	13.23 (8.66)	b = 0.121 (p = .801)	0.01
Sample Size	328	346		

Source: Six-month follow-up surveys administered six months after randomization

Notes: Analyses are linear mixed effects models with individuals nested within couple and include covariates. Effect size is Hedges' g, calculated by dividing the differences in estimated marginal means by the weighted pooled standard deviation (converted from model standard error).

Table V.2.d Phase I - Post-intervention estimated effects using data from 6-month follow-up — multilevel models

	ePREP vs. WL		OR vs. WL		ePREP vs. OR	
	B	SE	B	SE	B	SE
Intercept	12.01***	0.23	12.00***	0.23	13.30***	0.24
Intervention	2.42***	0.45	2.32***	0.45	0.12	0.48
Hispanic/Latino	0.70	0.52	0.16	0.50	0.08	0.56
White/Caucasian	-1.28***	0.38	-0.92*	0.38	-1.25**	0.41
Baseline Marital Status	-0.70	0.46	-0.11	0.45	-1.35**	0.48

Source: Baseline surveys and 6-month follow-up surveys.

Notes: Significant results are noted with asterisks: * p < .05, ** p < .01, *** p < .001. In these multilevel models, "intercept" coefficients are the estimated means for the comparison group, and the coefficients reflect the estimated difference from those means for the intervention group.

Mean imputation was not possible for the CSI-4 (relationship satisfaction) because the items are answered on different response scales, thus only complete cases will be used for this measure. It is also unnecessary for both partners to agree on their level of relationship satisfaction; therefore, although surveys are administered separately to each partner, we did not reconcile their responses. As noted above, in order to account for attrition and missing data, the analytic sample presented in this report is based only on the subsample of couples where both partners have completed the CSI-4 at baseline and the 6-month follow-up.

For Phase I, data for relationship satisfaction was analyzed using linear mixed effects models with individuals nested within couples. Six-month relationship satisfaction was the outcome with baseline relationship satisfaction entered at the individual level and intervention condition entered at the couple level as predictors. Three sets of analyses were run (ePREP vs. Waitlist, OurRelationship vs. Waitlist, ePREP vs. OurRelationship) and included race, ethnicity, and baseline marital status as covariates to account for baseline inequivalence. Statistical significance was established at the .05 level. Effect size was calculated using Hedges' g.

Phase I assessed whether participants assigned to the OurRelationship or ePREP groups had significantly higher relationship satisfaction six months after randomization compared to a Waitlist Control group. Six months after randomization, those assigned to either OurRelationship or ePREP had significantly higher relationship satisfaction than those assigned to the Waitlist Control group after controlling for race, ethnicity, and baseline marital status ($d_{OR}=.45, p<.001$; $d_{ePREP}=.47, p<.001$; see Tables V.1.a & b). In addition, Phase I assessed whether participants assigned to the OurRelationship and ePREP groups had significantly different levels of relationship satisfaction six months after randomization compared to one another. Results demonstrate that the OurRelationship and ePREP groups did not significantly differ in their relationship satisfaction six months after randomization ($p=.801$; see Table V.2.c) and actually had similarly sized impacts on relationship satisfaction (ePREP $d=.47$, OurRelationship $d=.45$), suggesting that both programs are effective in improving relationship satisfaction at follow-up compared to a Waitlist Control group. See Table V.2.d for the estimated effects of OurRelationship and ePREP accounting for covariates.

Key Findings – Phase II:

Phase II results showed that there were no significant differences in relationship satisfaction six months after randomization between ePREP and OurRelationship No Coach or Contingent Coach compared to the Coach condition. However, couples assigned to OurRelationship or ePREP No Coach or Contingent Coach demonstrated significantly higher relationship satisfaction six months after randomization compared to a Waitlist Control group.

Table V.2.d. Covariates included in Phase II impact analyses

Covariate	Description of the covariate
Gender	Gender is Female (1) or Male (0)
Hispanic/Latino	Ethnicity is Hispanic or Latino (1) versus not Hispanic or Latino (0)
Caucasian/White	Race is Caucasian/White (1) versus not Caucasian/White (0)
Marital Status	Couple was married (1) versus unmarried (0) at baseline
Relationship Satisfaction	Baseline relationship satisfaction (sum of CSI-4)

Notes: Based on analyses of baseline equivalence, not all Phase II impact analyses include all covariates. Refer to tables of Post-Intervention Estimate Effects (Tables V.2.m & n) for specific information on which covariates were included in each analysis.

Table V.2.e. Post-intervention estimated effects for OurRelationship No Coach vs. Coach using data from 6-month follow-up to address the primary research questions

Outcome measure	OR No Coach Mean or % (SD)	OR Coach Mean or % (SD)	OR No Coach versus OR Coach regression coefficient (p-value of difference)	Effect size
Relationship Satisfaction (6-month follow-up)	13.79 (6.63)	13.88 (8.64)	$b = 0.092$ ($p = .863$)	0x.01
Sample Size	160	262	-	-

Source: Six-month follow-up surveys administered six months after randomization

Notes: Analyses are linear mixed effects models with individuals nested within couple and include covariates. Effect size is Hedges' g, calculated by dividing the differences in estimated marginal means by the weighted pooled standard deviation (converted from model standard error).

Table V.2.f. Post-intervention estimated effects for OurRelationship Contingent Coach vs. Coach using data from 6-month follow-up to address the primary research questions

Outcome measure	OR Contingent Coach Mean or % (SD)	OR Coach Mean or % (SD)	OR Contingent Coach versus OR Coach regression coefficient (p-value of difference)	Effect size
Relationship Satisfaction	13.79 (6.75)	14.00 (8.62)	$b = 0.182$ ($p = .768$)	0.03
Sample Size	160	262	-	-

Source: Six-month follow-up surveys administered six months after randomization

Notes: Analyses are linear mixed effects models with individuals nested within couple and include covariates. Effect size is Hedges' g, calculated by dividing the differences in estimated marginal means by the weighted pooled standard deviation (converted from model standard error).

Table V.2.g. Post-intervention estimated effects for ePREP No Coach vs. Coach using data from 6-month follow-up to address the primary research questions

Outcome measure	ePREP No Coach Mean or % (SD)	ePREP Coach Mean or % (SD)	ePREP No Coach versus ePREP Coach regression coefficient (p-value of difference)	Effect size
Relationship Satisfaction	13.17 (5.20)	13.50 (5.40)	$b = -0.355$ ($p = .530$)	-0.04
Sample Size	184	246	-	-

Source: Six-month follow-up surveys administered six months after randomization

Notes: Analyses are linear mixed effects models with individuals nested within couple and include covariates. Effect size is Hedges' g, calculated by dividing the differences in estimated marginal means by the weighted pooled standard deviation (converted from model standard error).

Table V.2.h. Post-intervention estimated effects for ePREP Contingent vs. Coach using data from 6-month follow-up to address the primary research questions

Outcome measure	ePREP Contingent Coach Mean or % (SD)	ePREP Coach Mean or % (SD)	ePREP Contingent Coach versus ePREP Coach regression coefficient (p-value of difference)	Effect size
Relationship Satisfaction	13.43 (6.92)	13.45 (8.63)	$b = 0.016$ ($p = .977$)	0.00
Sample Size	158	246	-	-

Source: Six-month follow-up surveys administered six months after randomization

Notes: Analyses are linear mixed effects models with individuals nested within couple and include covariates. Effect size is Hedges' g , calculated by dividing the differences in estimated marginal means by the weighted pooled standard deviation (converted from model standard error).

Table V.2.i. Post-intervention estimated effects for OurRelationship No Coach vs. Waitlist Control using data from 6-month follow-up to address the primary research questions

Outcome measure	OR No Coach Mean or % (SD)	OR Waitlist Control Mean or % (SD)	OR No Coach versus Waitlist Control regression coefficient (p-value of difference)	Effect size
Relationship Satisfaction	13.82 (6.62)	11.49 (6.14)	$b = 2.324$ ($p < .001$)	0.36
Sample Size	160	160	-	-

Source: Six-month follow-up surveys administered six months after randomization

Notes: Analyses were conducted using the full OR No Coach group and a subsample of the Waitlist group matched on the above baseline characteristics using propensity score matching in the MatchIt R program. Analyses are linear mixed effects models with individuals nested within couple and include covariates. Effect size is Hedges' g , calculated by dividing the differences in estimated marginal means by the weighted pooled standard deviation (converted from model standard error).

Table V.2.j. Post-intervention estimated effects for OurRelationship Contingent Coach vs. Waitlist Control using data from 6-month follow-up to address the primary research questions

Outcome measure	OR Contingent Coach <i>Mean or % (SD)</i>	OR Waitlist Control <i>Mean or % (SD)</i>	OR Contingent Coach versus Waitlist Control regression coefficient (p-value of difference)	Effect size
Relationship Satisfaction	13.85 (10.55)	11.28 (7.41)	$b = 2.565$ ($p < .001$)	0.28
Sample Size	160	324	-	-

Source: Six-month follow-up surveys administered six months after randomization

Notes: Analyses are linear mixed effects models with individuals nested within couple and include covariates. Effect size is Hedges' g, calculated by dividing the differences in estimated marginal means by the weighted pooled standard deviation (converted from model standard error).

Table V.2.k. Post-intervention estimated effects for ePREP No Coach vs. Waitlist Control using data from 6-month follow-up to address the primary research questions

Outcome measure	ePREP No Coach <i>Mean or % (SD)</i>	ePREP Waitlist Control <i>Mean or % (SD)</i>	ePREP No Coach versus Waitlist Control regression coefficient (p-value of difference)	Effect size
Relationship Satisfaction	13.02 (8.37)	11.39 (6.31)	$b = 1.638$ ($p < .001$)	0.22
Sample Size	184	324	-	-

Source: Six-month follow-up surveys administered six months after randomization

Notes: Analyses are linear mixed effects models with individuals nested within couple and include covariates. Effect size is Hedges' g, calculated by dividing the differences in estimated marginal means by the weighted pooled standard deviation (converted from model standard error).

Table V.2.I. Post-intervention estimated effects for ePREP Contingent Coach vs. Waitlist Control using data from 6-month follow-up to address the primary research questions

Outcome measure	ePREP Contingent Coach <i>Mean or % (SD)</i>	ePREP Waitlist Control <i>Mean or % (SD)</i>	ePREP Contingent Coach versus Waitlist Control regression coefficient (p-value of difference)	Effect size
Relationship Satisfaction	13.21 (11.26)	11.29 (7.87)	$b = 1.914$ ($p = .003$)	0.20
Sample Size	158	324	-	-

Source: Six-month follow-up surveys administered six months after randomization

Notes: Analyses are linear mixed effects models with individuals nested within couple and include covariates. Effect size is Hedges' g , calculated by dividing the differences in estimated marginal means by the weighted pooled standard deviation (converted from model standard error).

Table V.2.m Post-intervention estimated effects for OurRelationship multilevel models using data from 6-month follow-up

	OR No Coach vs. OR Coach		OR Contingent vs. OR Coach		OR No Coach vs. WL		OR Contingent vs. WL	
	B	SE	B	SE	B	SE	B	SE
Intercept	13.83***	0.25	13.80***	0.30	12.26***	0.27	12.13***	0.28
Intervention	0.09	0.54	0.18	0.62	2.32***	0.49	2.56***	0.59
Gender	-0.01	0.33	-	-	0.52	0.40	-1.07	0.30
Hispanic/ Latino	0.23	0.58	-	-	0.59	0.63	-0.52	0.59
White/ Caucasian	-0.02	0.48	0.12	0.55	-0.15	0.48	-0.09	0.51
Marital Status at Baseline	0.04	0.52	-0.54	0.60	-0.28	0.52	-1.64**	0.55
CSI at Baseline	0.46	0.05***	-	-	0.58***	0.05	-	-

Source: Baseline surveys and 6-month follow-up surveys.

Notes: Significant results are noted with asterisks: * $p < .05$, ** $p < .01$, *** $p < .001$. In these multilevel models, "intercept" coefficients are the estimated means for the comparison group, and the coefficients reflect the estimated difference from those means for the intervention group.

"-" indicates that the variable was not included as a covariate in the model.

Table V.2.n Post-intervention estimated effects for ePREP multilevel models using data from 6-month follow-up

	ePREP No Coach vs. ePREP Coach		ePREP Contingent vs. ePREP Coach		ePREP No Coach vs. WL		ePREP Contingent vs. WL	
	B	SE	B	SE	B	SE	B	SE
Intercept	13.27***	0.28	13.44***	0.27	11.98***	0.22	11.92***	0.29
Intervention	-0.36	0.57	0.02	0.55	1.64***	0.47	1.91**	0.63
Gender	0.46	0.34	0.29	0.33	0.31	0.29	-1.19	0.30
Hispanic/ Latino	-0.22	0.63	0.14	0.60	-0.64	0.51	-0.30	0.58
White/ Caucasian	-1.06*	0.52	-	-	-0.54	0.43	-	-

	ePREP No Coach vs. ePREP Coach		ePREP Contingent vs. ePREP Coach		ePREP No Coach vs. WL		ePREP Contingent vs. WL	
Marital Status at Baseline	1.10	0.56	0.61	0.54	0.09	0.46	-	-
CSI at Baseline	0.52***	0.05	0.54***	0.05	0.60***	0.04	-	-

Source: Baseline surveys and 6-month follow-up surveys.

Notes: Significant results are noted with asterisks: * $p < .05$, ** $p < .01$, *** $p < .001$. In these multilevel models, “intercept” coefficients are the estimated means for the comparison group, and the coefficients reflect the estimated difference from those means for the intervention group.

“-” indicates that the variable was not included as a covariate in the model.

For Phase II, data for relationship satisfaction was analyzed using linear mixed effects models with individuals nested within couples. Six-month relationship satisfaction was the outcome with baseline relationship satisfaction entered at the individual level and intervention condition entered at the couple level as predictors. Eight sets of analyses were run (OurRelationship/ePREP No Coach vs. Coach, OurRelationship/ePREP Contingent Coach vs. Coach, OurRelationship/ePREP No Coach vs. Waitlist Control, OurRelationship/ePREP Contingent Coach vs. Waitlist Control) and included gender, race, ethnicity, marital status, and/or baseline relationship satisfaction as covariates when there was baseline inequivalence (see Tables V.1.d-k). Effect size was calculated using Hedges’ g . Statistical significance was established at the .05 level. In addition, given that the comparison of the OurRelationship No Coach and Waitlist Control groups revealed high differential attrition and differences in baseline marital status outside of the normal range, we used a propensity score matching procedure to select equivalent samples in each group that were maximally comparable. We used the R program MatchIt to select a subset of the Waitlist Control group that was maximally similar to the OurRelationship No Coach group in terms of baseline relationship satisfaction, ethnicity, race, gender, and marital status. This procedure ensured that the only meaningful difference between the OurRelationship No Coach and Waitlist Control groups was the treatment status and not baseline levels of the covariates (see Section B of the Appendix). See Tables V.2.m&n for the estimated effects of the Phase II intervention groups accounting for covariates.

Phase II assessed whether couples assigned to ePREP and OurRelationship No Coach or Contingent Coach had significantly higher relationship satisfaction six months after randomization compared to couples assigned to ePREP and OurRelationship Coach. Results showed that couples assigned to ePREP and OurRelationship No Coach did not have significantly higher relationship satisfaction six months after randomization compared to couples assigned to ePREP Coach ($d_{\text{ePREP}}=.04, p=.530$) or OurRelationship Coach ($d_{\text{OR}}=.01, p=.863$), suggesting similar program benefits for couples receiving automated email reminders (“No Coach”) and Coach contact (see Tables V.2.e&g). In addition, non-responsive ePREP and OurRelationship couples who were re-randomized to receive a Coach (“Contingent Coach”) did not have significantly higher relationship satisfaction six months after randomization compared to couples assigned to ePREP Coach ($d_{\text{ePREP}}=.00, p=.977$) or OurRelationship Coach ($d_{\text{OR}}=.03$,

$p=.768$), suggesting similar benefits for couples who receive coaching later in the program and couples who receive coaching starting at enrollment (see Tables V.2.f&h).

Phase II also assessed whether couples assigned to ePREP and OurRelationship No Coach or Contingent Coach had significantly higher relationship satisfaction six months after randomization compared to a Waitlist Control group. Results showed that couples assigned to OurRelationship or ePREP No Coach had significantly higher relationship satisfaction six months after randomization compared to couples assigned to the Waitlist Control group ($d_{OR}=.36, p<.001$; $d_{ePREP}=.22, p<.001$). In addition, couples assigned to ePREP and OurRelationship Contingent Coach had significantly higher relationship satisfaction six months after randomization compared to couples assigned to the Waitlist Control group ($d_{OR}=.28, p<.001$; $d_{ePREP}=.20, p=.003$). Taken together, these findings suggest that engaging in relationship education is a useful approach for improving relationship satisfaction, regardless of the amount of coach contact (see Tables V.2.i-l).

VI. Discussion

Phase I

Findings from the implementation evaluation for Phase I presented in this report demonstrate that online relationship interventions are a viable means of reaching low-income couples. Couples found these interventions online, enrolled in them, completed them, and, as the impact analyses show, benefitted from the programs offered. These kinds of interventions can reach many more couples than can be served in in-person services. The results of the impact evaluations conducted as part of this report further demonstrate their promise in improving relationships.

Results from Phase I of the impact evaluation show that being assigned to either ePREP or OurRelationship led to higher relationship satisfaction compared to Waitlist Control group, as measured six months later. Notably, the effect sizes on relationship satisfaction in the current study and Doss et al. (2020), a paper published with the Phase I data, are larger than previous studies that have investigated the effectiveness of curricula developed specifically for low-income couples (Barton et al., 2018).

There are likely several contributing factors to the findings that ePREP and OurRelationship worked to increase relationship satisfaction. First, these are programs based on in-person interventions that have each shown evidence of effectiveness in rigorous studies. Second, the study aided couples who were already seeking help for their relationships (primarily through online searches using help-seeking terms) rather than recruiting couples from the community to attend. As a result, couples in the current program were likely more motivated and receptive to implementing skills that would help to improve relationship satisfaction. Finally, both online programs were self-directed and required engagement from both partners, which may have contributed to greater learning and application of concepts by couples; helping to foster improved relationship satisfaction. Indeed, the implementation findings demonstrated that 68% of couples assigned to OurRelationship and 69% of couples assigned to ePREP completed all program components. In addition, overall ratings of OurRelationship and ePREP, as well as staff-participant interactions were rated positively, suggesting overall participant engagement in both intervention groups.

In addition, in the current report, there were no significant differences in relationship satisfaction six months after randomization between OurRelationship and ePREP. These findings suggest that both web-based relationship education programs are effective in improving relationship satisfaction six months after enrollment. Indeed, we would expect few differences in the two online programs' impact on relationships, as both are designed to strengthen couples' relationships and have been shown to improve functioning in several relationship domains (Braithwaite & Fincham, 2014; Doss et al., 2016).

Phase I findings suggest that these online, evidence-based relationship interventions are effective. Future research should continue to examine new adaptations and components of these interventions and test them among different types of couples or relationship issues.

Phase II

Phase II of this impact evaluation extended the work of Phase I by evaluating varying types of coach contact on program effects. Results from Phase II show that couples assigned to ePREP and OurRelationship No Coach did not have significantly higher relationship satisfaction six months after randomization compared to couples assigned to ePREP and OurRelationship Coach. In addition, couples assigned to ePREP and OurRelationship Contingent Coach did not have significantly higher relationship satisfaction six months after randomization compared to couples assigned to ePREP and OurRelationship Coach. Taken together, these findings suggest that the addition of coach contact in web-based intervention programming does not significantly impact program effects regarding relationship satisfaction, at least for this kind of population of help-seeking, low-income couples. This observation is important when considering the cost structure of web-based relationship education programs. The removal of coach contact could help programs save money without sacrificing treatment effects; further increasing the feasibility of implementing such programming and continuing to extend the reach of relationship education to couples in need. At the same time, implementation findings demonstrated that couples assigned to Coach had significantly higher program completion rates compared to couples assigned to Contingent Coach or No Coach. Thus, coach contact appears to play a pivotal role in program implementation by aiding couples in engaging with more of the program. Also note that we did not examine the impact of coaching on other indicators of relationship quality and success, e.g., stability or conflict.

These findings also corroborate Phase I findings, suggesting that couples assigned to complete web-based relationship education demonstrate higher levels of relationship satisfaction six months after randomization compared to couples in a no-treatment control group, regardless of the type of coach contact.

Taken together, both Phase I and Phase II demonstrate the utility of employing web-based relationship education programming for improving low-income help-seeking couples' relationships. In particular, there are a few key aspects of this type of programming that might make it particularly practical and useful for relationship education service delivery among low-income couples. First, as shown in the current report, online relationship education can be employed flexibly (e.g., with or without a coach) and with fidelity, making it a feasible alternative or addition to the repertoire of in-person relationship education programming currently available. Further, this type of online intervention is particularly well-suited for low-income populations, who tend to experience higher rates of relationship distress than the general population. Online interventions circumvent many of the barriers that low-income couples face in accessing relationship services, such as fees, transportation, and work conflicts, making them important options to ensure that low-income couples receive much-needed services. Additionally, the online recruitment and delivery of these interventions make them well-

positioned to attract and engage with couples at a time when they are seeking help and therefore are likely motivated to take part in relationship services. Finally, the general equivalence of the ePREP and OurRelationship programs indicates that both communication-focused and problem/insight-focused interventions can be successful in improving the relationship satisfaction of low-income couples, thus providing two comparable intervention options for the field. Future work should continue to explore implementation strategies and impacts of web-based relationship education programming on individual and relationship outcomes, as well as moderators of these effects.

VII. References

- Barton, A. W., Beach, S. R. H., Wells, A. C., Ingels, J. B., Corso, P. S., Sperr, M. C., Anderson, T. N., & Brody, G. H. (2018). The protecting strong african american families program: A randomized controlled trial with rural african american couples. *Prevention Science, 19*(7), 904–913. <https://doi.org/10.1007/s11121-018-0895-4>
- Braithwaite, S. R., & Fincham, F. D. (2011). Computer-based dissemination: A randomized clinical trial of ePREP using the actor partner interdependence model. *Behaviour Research and Therapy, 49*(2), 126–131. <https://doi.org/10.1016/j.brat.2010.11.002>
- Bramlett, M. D., & Mosher, W. D. (2002). Cohabitation, marriage, divorce, and remarriage in the United States [Data set]. In *Vital Health Stat* (Vol. 23, Issue 22). <https://doi.org/10.1037/e372462004-001>
- Cunradi, C. B., Caetano, R., & Schafer, J. (2002). Socioeconomic predictors of intimate partner violence among white, black, and hispanic couples in the United States. *Journal of Family Violence, 13*.
- Doss, B. D. (2014). Oklahoma Marriage Initiative, Baseline Report. *Unpublished Analyses*.
- Doss, B. D., Cicila, L. N., Georgia, E. J., Roddy, M. K., Nowlan, K. M., Benson, L. A., & Christensen, A. (2016). A randomized controlled trial of the web-based OurRelationship program: Effects on relationship and individual functioning. *Journal of Consulting and Clinical Psychology, 84*(4), 285–296. <https://doi.org/10.1037/ccp0000063>
- Doss, B. D., Cicila, L. N., Hsueh, A. C., Morrison, K. R., & Carhart, K. (2014). A randomized controlled trial of brief coparenting and relationship interventions during the transition to parenthood. *Journal of Family Psychology, 28*(4), 483–494. <https://doi.org/10.1037/a0037311>
- Doss, B. D., Knopp, K., Roddy, M. K., Rothman, K., Hatch, S. G., & Rhoades, G. K. (2020). Online programs improve relationship functioning for distressed low-income couples: Results from a nationwide randomized controlled trial. *Journal of Consulting and Clinical Psychology, 88*(4), 283–294. <https://doi.org/10.1037/ccp0000479>
- Lei, H., Nahum-Shani, I., Lynch, K., Oslin, D., & Murphy, S. A. (2012). A "SMART" design for building individualized treatment sequences. *Annual Review of Clinical Psychology, 8*, 14.1–14.28. doi: 10.1146/annurev-clinpsy-032511-143152
- Lundquist, E., Hsueh, J., Lowenstein, A. E., Faucetta, K., Gubits, D., Michalopoulos, C., & Knox, V. (2014). A family-strengthening program for low-income families: Final impacts from the Supporting Healthy Marriage evaluation. *New York, NY: MDRC*.

VIII. Appendices

A. Data and study sample

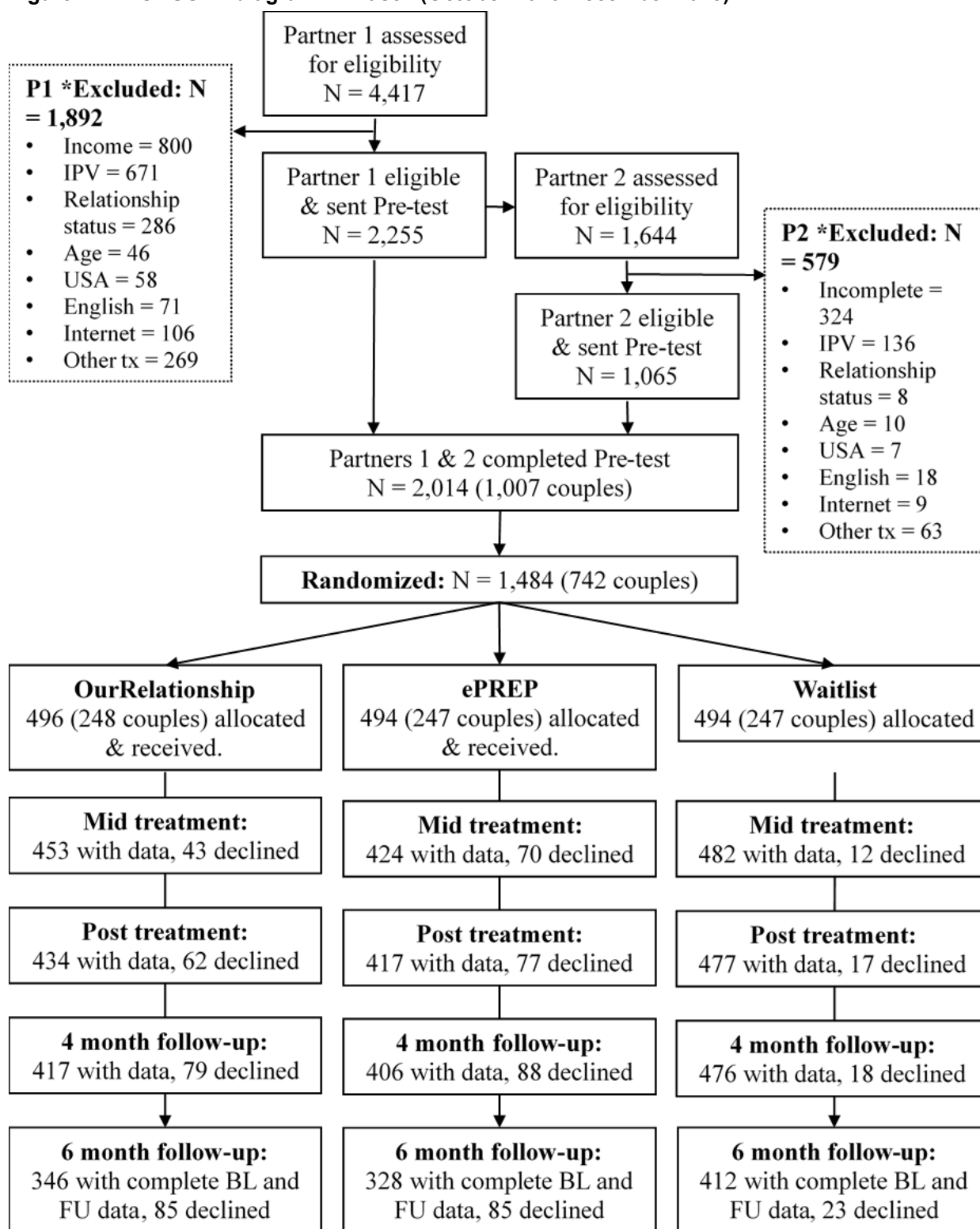
Table A.1. Data used to address implementation research questions

Implementation element	Research question	Data source	Timing/ frequency of data collection	Party responsible for data collection
Intervention Group				
Fidelity	Were all intended workshop components offered and for the expected duration?	Server logs	24-7 / 365	Website hosting company
Fidelity	Were there adaptations to the workshops over the course of the study?	Server logs and project tracking logs	Reviewed quarterly	Website hosting company; intervention staff
Fidelity	Were individual service contacts delivered reliably across providers to those in the intervention group?	Project tracking logs	Weekly	Intervention staff
Dosage	What percentage of couples who enrolled in the intervention completed 100% of the intervention (i.e., workshops, individual service contacts)?	Server logs; Project tracking logs	All sessions delivered	Study staff
Dosage	What percentage of couples who enrolled in the intervention completed 2/3 of the intervention (i.e., workshops, individual service contacts)?	Server logs; Project tracking logs	All sessions delivered	Study staff
Quality	What was the quality of staff—participant interactions during the intervention (i.e., workshops and individual service contacts)?	Three items tapping therapeutic alliance with the coach (Phase 1 only)	After each individual service contact	Study staff / local evaluator
Engagement	What was intervention couples' overall feedback on the program?	Two survey items on program satisfaction	At Post-Program Survey	Study staff / local evaluator
Context	What outside services were utilized by couples in the intervention group during the evaluation period?	Survey item on other services used during evaluation period	At Post-Program Survey	Study staff / local evaluator

Implementation element	Research question	Data source	Timing/ frequency of data collection	Party responsible for data collection
Counterfactual (Waitlist Control Group)				
Fidelity	Was the initial Welcome call offered to all couples in the control/comparison group and for the expected duration?	Project tracking logs	All initial calls with control couples	Study and intervention staff
Fidelity	What outside services were utilized by couples in the control group during the evaluation period?	Survey item on other services used during evaluation period	At 2 months, 4 months and 6 months post randomization. Note: This question was not asked in Phase II at 2 months post randomization.	Study staff / local evaluator

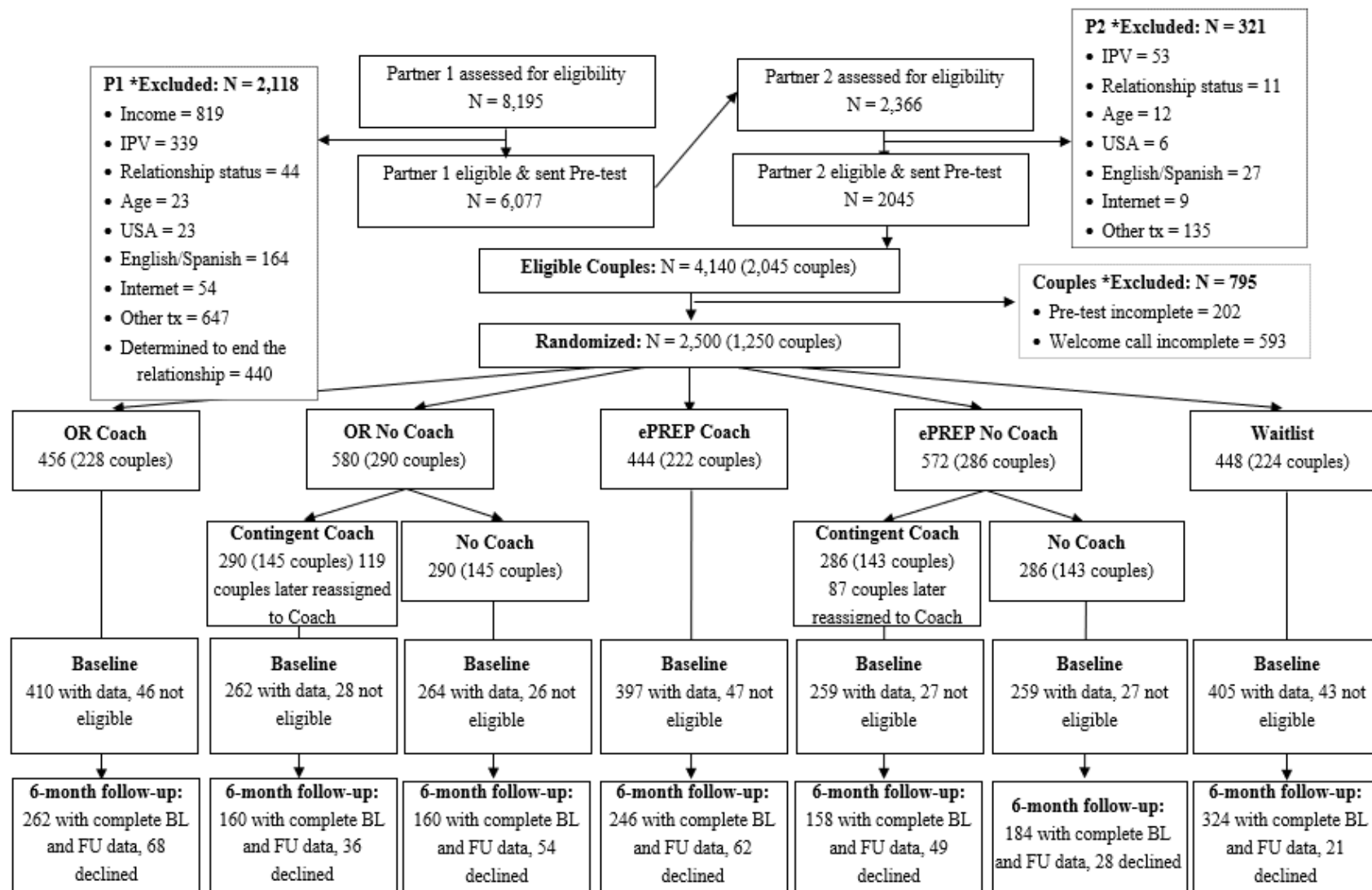
Table A.2. Key features of the impact analysis data collection

	Data source	Timing of data collection	Mode of data collection	Party responsible for data collection	Start and end date of data collection
Intervention	Phase I: ePREP or OR Phase II: ePREP Coach/No Coach, OR Coach/No Coach	Enrollment (<u>Baseline</u>); 1 month after randomization (<u>Mid</u>); 2 months after randomization (<u>Post</u>); 4 months after randomization (<u>4-month follow-up</u>); 6 months after randomization (<u>6-month follow-up</u>)	Online survey (Qualtrics and nFORM)	Program staff (via monitoring online data collection)	Phase I: October 2016-December 2018 Phase II: January 2019-June 2020
Counterfactual (Waitlist Control)	Comparison group study participants	Enrollment (<u>Baseline</u>); 1 month after randomization (<u>Mid</u>); 2 months after randomization (<u>Post</u>); 4 months after randomization (<u>4-month follow-up</u>); 6 months after randomization (<u>6-month follow-up</u>)	Online survey (Qualtrics and nFORM)	Program staff (via monitoring online data collection)	Phase I: October 2016-December 2018 Phase II: January 2019-June 2020

Figure A.1. CONSORT diagram – Phase I (October 2016-December 2018)

Note: "Complete" data consists of cases wherein both partners within the same couple completed the CSI-4.

Figure A.2. CONSORT diagram – Phase II (January 2019-May 2020)



Note: “Not eligible” indicates participants who had not yet reached their 6-month follow-up at the writing of this report and therefore did not contribute to the impact report findings. “Complete” 6-month follow-up data consists of cases wherein both partners within the same couple completed the CSI-4.

B. Baseline equivalence summary findings and matching approach in Phase II

Table B.1. Phase II - Summary statistics of key baseline measures and baseline equivalence across study groups (OurRelationship No Coach and OurRelationship Coach), for couples completing relationship satisfaction measure (CSI-4) at 6-month follow-up

Baseline measure	OR No Coach versus OR Coach regression coefficient			
	OR No Coach Mean (SD) or %	OR Coach Mean (SD) or %	(p-value)	Effect size
Female (%)	57%	52%	b = 0.185 (p = .359)	0.11
Race/ethnicity (%)				
Hispanic/Latino	17%	13%	b = 0.374 (p = .667)	0.17
Caucasian/White	66%	72%	b = -0.560 (p = .535)	-0.18
Married (%)	61%	42%	b = 0.766 (p < .001)	0.46
Relationship satisfaction (range: 0 to 21)	10.93 (4.98)	10.24 (5.35)	b = 0.688 (p = .297)	0.13
Sample size	160	262	n.a.	n.a.

Notes: p-values are included in parentheses. Effect sizes for continuous variables are Hedges' g, calculated by dividing the differences in means by the weighted pooled standard deviation. Effect sizes for dichotomous variables are Cox indices in comparable units.

n.a. = not applicable.

Table B.2. Phase II - Summary statistics of key baseline measures and baseline equivalence across study groups (OurRelationship Contingent Coach and OurRelationship Coach), for couples completing relationship satisfaction measure (CSI-4) at 6-month follow-up

Baseline measure	OR Contingent Coach <i>Mean (SD) or %</i>	OR Coach <i>Mean (SD) or %</i>	OR Contingent Coach versus OR Coach regression coefficient (p-value)	Effect size
Female (%)	52%	52%	b = 0.008 (p = .967)	0.01
Race/ethnicity (%)				
Hispanic/Latino	12%	13%	b = -0.017 (p = .984)	-0.01
Caucasian/White	69%	72%	b = -0.165 (p = .866)	-0.08
Married (%)	44%	42%	b = 0.082 (p = .686)	0.05
Relationship satisfaction (range: 0 to 21)	10.09 (4.93)	10.24 (5.35)	b = -0.149 (p = .819)	-0.03
Sample size	160	262	n.a.	n.a.

Notes: p-values are included in parentheses. Effect sizes for continuous variables are Hedges' g, calculated by dividing the differences in means by the weighted pooled standard deviation. Effect sizes for dichotomous variables are Cox indices in comparable units.

n.a. = not applicable.

Table B.3. Phase II - Summary statistics of key baseline measures and baseline equivalence across study groups (ePREP No Coach and ePREP Coach), for couples completing relationship satisfaction measure (CSI-4) at 6-month follow-up

Baseline measure	ePREP No Coach <i>Mean (SD) or %</i>	ePREP Coach <i>Mean (SD) or %</i>	ePREP No Coach versus ePREP Coach regression coefficient (p-value)	Effect size
Female (%)	50%	54%	$b = -0.179$ ($p = .358$)	-0.11
Race/ethnicity (%)				
Hispanic/Latino	17%	12%	$b = 0.482$ ($p = .573$)	0.25
Caucasian/White	70%	73%	$b = -0.293$ ($p = .634$)	-0.08
Married (%)	45%	52%	$b = -0.278$ ($p = .156$)	-0.17
Relationship satisfaction (range: 0 to 21)	10.58 (4.82)	10.36 (4.74)	$b = 0.220$ ($p = .696$)	0.05
Sample size	184	246	n.a.	n.a.

Notes: p-values are included in parentheses. Effect sizes for continuous variables are Hedges' g, calculated by dividing the differences in means by the weighted pooled standard deviation. Effect sizes for dichotomous variables are Cox indices in comparable units.

n.a. = not applicable.

Table B.4. Phase II - Summary statistics of key baseline measures and baseline equivalence across study groups (ePREP Contingent Coach and ePREP Coach), for couples completing relationship satisfaction (CSI-4) at 6-month follow-up

Baseline measure	ePREP Contingent Coach <i>Mean (SD) or %</i>	ePREP Coach <i>Mean (SD) or %</i>	ePREP Contingent Coach versus ePREP Coach regression coefficient (p-value)	Effect size
Female (%)	53%	54%	$b = -0.078$ ($p = .703$)	-0.05
Race/ethnicity (%)				
Hispanic/Latino	16%	12%	$b = 0.341$ ($p = .246$)	0.21
Caucasian/White	73%	73%	$b = 0.109$ ($p = .898$)	0.02
Married (%)	47%	52%	$b = -0.208$ ($p = .308$)	-0.13
Relationship satisfaction (range: 0 to 21)	9.87 (5.29)	10.36 (4.74)	$b = -0.488$ ($p = .432$)	0.10
Sample size	158	246	n.a.	n.a.

Notes: p-values are included in parentheses. Effect sizes for continuous variables are Hedges' g, calculated by dividing the differences in means by the weighted pooled standard deviation. Effect sizes for dichotomous variables are Cox indices in comparable units.

n.a. = not applicable.

Table B.5. Phase II - Summary statistics of key baseline measures and baseline equivalence across study groups (OurRelationship No Coach and waitlist control), for couples completing relationship satisfaction measure (CSI-4) at 6-month follow-up

Baseline measure	OR No Coach Mean (SD) or %	OR Waitlist Control Mean (SD) or %	OR No Coach versus Waitlist Control regression coefficient (p-value)	Effect size
Female (%)	57%	55%	$b = 0.066$ ($p = .734$)	0.04
Race/ethnicity (%)				
Hispanic/Latino	17%	14%	$b = 0.277$ ($p = .731$)	0.12
Caucasian/White	66%	73%	$b = -0.657$ ($p = .437$)	-0.21
Married (%)	61%	48%	$b = 0.557$ ($p = .005$)	0.34
Relationship satisfaction (range: 0 to 21)	10.93 (4.98)	10.10 (5.14)	$b = 0.820$ ($p = .179$)	0.16
Sample size	160	324	n.a.	n.a.

n.a. = not applicable.

Notes: p-values are included in parentheses. Effect sizes for continuous variables are Hedges' g, calculated by dividing the differences in means by the weighted pooled standard deviation. Effect sizes for dichotomous variables are Cox indices in comparable units.

Table B.6. Phase II - Summary statistics of key baseline measures and baseline equivalence across study groups (OurRelationship Contingent Coach and waitlist control), for couples completing relationship satisfaction measure (CSI-4) at 6-month follow-up

Baseline measure	OR Contingent Coach Mean (SD) or %	OR Waitlist Control Mean (SD) or %	OR Contingent Coach versus Waitlist Control regression coefficient (p-value)	Effect size
Female (%)	52%	55%	$b = -0.111$ ($p = .563$)	-0.07
Race/ethnicity (%)				
Hispanic/Latino	12%	14%	$b = -0.088$ ($p = .911$)	-0.09
Caucasian/White	69%	73%	$b = -0.183$ ($p = .833$)	-0.10
Married (%)	44%	48%	$b = -0.127$ ($p = .513$)	-0.08
Relationship satisfaction (range: 0 to 21)	10.09 (4.91)	10.10 (5.14)	$b = -0.017$ ($p = .977$)	0.00
Sample size	160	324	n.a.	n.a.

Notes: p-values are included in parentheses. Effect sizes for continuous variables are Hedges' g, calculated by dividing the differences in means by the weighted pooled standard deviation. Effect sizes for dichotomous variables are Cox indices in comparable units.

n.a. = not applicable.

Table B.7. Phase II - Summary statistics of key baseline measures and baseline equivalence across study groups (ePREP No Coach and waitlist control), for couples completing relationship satisfaction measure(CSI-4) at 6-month follow-up

Baseline measure	ePREP No Coach Mean (SD) or %	Waitlist Control Mean (SD) or %	ePREP No Coach versus Waitlist Control regression coefficient (p-value)	Effect size
Female (%)	50%	55%	$b = -0.211$ ($p = .255$)	-0.13
Race/ethnicity (%)				
Hispanic/Latino	17%	14%	$b = 0.240$ ($p = .763$)	0.12
Caucasian/White	70%	73%	$b = -0.219$ ($p = .781$)	-0.08
Married (%)	45%	48%	$b = -0.097$ ($p = .599$)	-0.06
Relationship satisfaction (range: 0 to 21)	10.58 (4.82)	10.10 (5.14)	$b = 0.477$ ($p = .402$)	0.10
Sample size	184	324	n.a.	n.a.

Notes: p-values are included in parentheses. Effect sizes for continuous variables are Hedges' g, calculated by dividing the differences in means by the weighted pooled standard deviation. Effect sizes for dichotomous variables are Cox indices in comparable units.

n.a. = not applicable.

Table B.8. Phase II - Summary statistics of key baseline measures and baseline equivalence across study groups (ePREP Contingent Coach and waitlist control), for couples completing relationship satisfaction measure (CSI-4) at 6-month follow-up

Baseline measure	ePREP Contingent Coach <i>Mean (SD) or %</i>	Waitlist Control <i>Mean (SD) or %</i>	ePREP Contingent Coach versus Waitlist Control regression coefficient (p-value)	Effect size
Female (%)	53%	55%	$b = -0.109$ ($p = .574$)	-0.07
Race/ethnicity (%)				
Hispanic/Latino	16%	14%	$b = 0.192$ ($p = .811$)	0.07
Caucasian/White	73%	73%	$b = 0.062$ ($p = .942$)	0.02
Married (%)	47%	48%	$b = 0.028$ ($p = .886$)	0.02
Relationship satisfaction (range: 0 to 21)	9.87 (5.29)	10.10 (5.14)	$b = -0.232$ ($p = .711$)	-0.04
Sample size	158	324	n.a.	n.a.

Notes: p-values are included in parentheses. Effect sizes for continuous variables are Hedges' g, calculated by dividing the differences in means by the weighted pooled standard deviation. Effect sizes for dichotomous variables are Cox indices in comparable units.

n.a. = not applicable.

Baseline equivalence tests indicated that the OurRelationship No Coach and Waitlist Control groups had different proportions of married participants, and this difference was larger than the acceptable range given our attrition rates according to WWC evidence review standards. Therefore, we used a propensity score matching procedure to rebalance the samples for this comparison. We used the MatchIt program in R to select a subsample of the waitlist group matched to the OurRelationship No Coach group on the baseline covariates (marital status, gender, ethnicity, race, and baseline relationship satisfaction). Although this procedure sacrifices the randomization of the RCT, MatchIt uses propensity score weighting to select equivalent samples that are maximally comparable, in order to ensure that the only meaningful difference between the treatment and control group is the treatment status and not baseline levels of the covariates. Nearest neighbor matching was used, which selects the best match from the control group (Waitlist Control) for each case in the treatment group (OurRelationship No Coach) in a way that minimizes the “distance” (i.e., maximizes fit to the propensity score model). See Tables B.4 and B.8 in the Appendix for more information on baseline equivalence among these groups. Balance of the matched samples improved, as indicated by a percent balance improvement of 94% and an average mean difference between Waitlist and OurRelationship No Coach groups across all variables of 0.04 for the matched samples, compared to 0.14 for the unmatched samples.

C. DATA PREPARATION

For both Phase I and Phase II, data were downloaded from Qualtrics into SPSS. A separate dataset was created and cleaned for each survey time point, such that each partner who completed the survey was accounted for at each time point and all measures were scored. Each dataset was then merged to create one stacked dataset with each time point represented for each participant. For item nonresponse when creating scales (and when missingness is 20 percent or less for the items needed to construct the scale), we typically use mean imputation. However, mean imputation is not possible for the CSI-4 (relationship satisfaction) because the items are answered on different response scales, thus only complete cases will be used for this measure. It is also unnecessary for both partners to agree on their level of relationship satisfaction; therefore, although surveys are administered separately to each partner, we did not reconcile their responses.

D. IMPACT ESTIMATION

The baseline equivalence were linear mixed-effects models that predicted each baseline covariate of interest by each treatment group comparison. A random intercept for the couple identifier was estimated in order to account for nesting of participants within couples. The lmer function from the R package lme4 was used for these analyses. In lme4, the model was specified as:

$$\text{Covariate} \sim \text{Treatment} + (1|\text{Couple ID})$$

where Covariate is each baseline covariate, estimated separately (i.e., gender, race, ethnicity, marital status, and baseline relationship satisfaction); the tilde operator represents a linear regression; Treatment is each treatment group comparison (two comparisons for Phase 1 and eight comparisons for Phase 2, as described above), estimated separately; and Couple ID is the couple identifier, while the number 1 and the vertical bar preceding it represent that a random intercept is specified.

The program impact analyses took the same general form as the baseline equivalence models, linear mixed-effects models with individuals nested within couples. In these models, relationship satisfaction at the 6-month follow-up was predicted by each treatment group comparison, controlling for the baseline covariates that differed between the two treatment comparison groups to an extent greater than the minimum guideline specified by the WWC evidence review standards (i.e., $d > 0.05$), based on the previous baseline equivalence analyses. Thus, the specific covariates included in each impact analysis was different depending on the treatment groups being compared. In lme4, the model was specified as:

$$6\text{month Relationship Satisfaction} \sim \text{Treatment} [+ \text{Covariates}] + (1|\text{Couple ID})$$

E. ADDITIONAL TABLES

Table E.1. Description of ePREP Structure and Content

ePREP Program Structure and Content			
Coach Meeting 1: Welcome			
Section 1:	Welcome (background on program, rooted in research, The Three Keys, Communication Danger Signs, Time Out, Wrap up and Review	Homework Assignment 1: Plan Time Out strategy for their relationship	Section 1 Quiz
Section 2:	Speaker-Listener Technique (SLT), Wrap up and Review	Homework Assignment 2: Practice SLT together, brainstorm SLT topic for coach call	Section 2 Quiz
Coach Meeting 2: PREP2 Call			
Section 3:	It's a little thing, Beneath the Surface, Events, Issues, and Hidden Issues, Wrap up and Review	Homework Assignment 3: Practice SLT with Issues and Hidden Issues	Section 3 Quiz
Section 4:	XYZ Statements, Problem Solving, Managing Expectations, Wrap up and Review	Homework Assignment 4: Implement Problem Solving experiment from presentation	Section 4 Quiz
Coach Meeting 3: PREP3 Call			
Section 5:	Understanding Commitment, Building Dedication, Wrap Up and Review	Homework Assignment 5: Discuss vision of relationship together	Section 5 Quiz
Section 6:	Having Fun Together, How to Keep the Fun Alive (talking as friends, talking supportively), Recap and Review of Sections 1-6, Wrap Up	Homework Assignment 6: Create and implement Fun Deck	Section 6 Quiz
Coach Meeting 4: PREP4 Call			

Table E.2. Description of ePREP Structure and Content

^aOurRelationship Program Structure and Content		
Observe Phase		
Chapter 1	Introduction	Introduction to OurRelationship.com
Chapter 2	A Snapshot of my Relationship	Individual completes assessment to assess relationship strengths and difficulties.
	My Results	Individual provided feedback based on assessment about their and their partner's relative strengths and difficulties compared to the average US couple.
Chapter 3	Selecting my Core Issue	Individual provided advice on selecting a core issue; user selects core issue. Individual picks a few fun activities they would like to do with their partner over the coming months and is introduced to the Speaker Listener structure (structure used in the Observe Conversation).
	Rating my Relationship	Individual completes an assessment to track progress.
Chapter 4	Observe Conversation	Couples work together to finalize a core issue or issues to work on during the program. Includes quick assessment about core issue(s) (drives feedback later in the program) and how conversation went.
Chapter 5	What's Next?	Individual is shown which fun activities their partner chose to do together over the coming months and is provided quick preview of Understand Phase.
Coach Meeting 2: Observe Call		
Understand Phase		
Chapter 6	Escaping the Blame Trap	Individuals are introduced to understanding their Differences, Emotional Reactions/Sensitivities, External Circumstances/Stressors, Pattern of Interaction (DEEP) and provided education on how they can stop blaming their partner.
Chapter 7	Differences	Education on how natural differences (e.g., personality) can create relationship problems. Individuals are provided feedback on their differences in personality traits, emotional expressiveness, and desire for closeness (based on earlier assessments). Individuals select the biggest difference in their relationship that impacts the core issue most. They then are asked to talk about the effect of this difference and any advantages.
	Emotions	Discussion of hidden and surface emotions. Individuals select the hidden and surface emotions that arise around their own core issue.
	External Stress	Discussion of how external stress impacts core issue. Individuals are asked to think about how stress impacts their own relationship.
	Patterns	Individuals learn about how patterns of communication impact core issue and how these patterns develop over time. Individuals identify the pattern they and their partner get into during a fight. They then identify the pattern they get into after a fight (including recovery method, recovery speed, and role of apologies).
	Putting the Pieces Together	Individuals consider how their relationship has changed over time and how the DEEP components they've identified have contributed.
Chapter 8	Partner Core Issue *only for 2 core issue couples*	Individuals select the DEEP components that may impact their partner's core issue.

Chapter 9	Speaking Skills	Individuals are reminded of the Speaker-Listener structure (structure used in Understand conversation) and are given tips on how to be effective in their role as speaker.
	Listening Skills	Individuals are given tips on how to be effective in their role as listener.
	Editing What You've Written	Individuals have ability to review and edit what they've written in Chapter 7 (DEEP components and patterns) which will be displayed during the Understand Conversation.
Chapter 10	Understand Conversation- 1 Core Issue	Couples complete their Understand conversation wherein they share their DEEP understanding with one another. Participants complete quick assessment on how the conversation went.
	Understand Conversation- 2 Core Issues- First	If couples have two core issues, couples complete their first Understand conversation wherein they share their DEEP understanding with one another of one of the partner's core issues. Participants complete quick assessment on how the conversation went.
	Understand Conversation- 2 Core Issues- Second	If couples have two core issues, couples complete their second Understand conversation wherein they share their DEEP understanding with one another of the second partner's core issue. Participants complete quick assessment on how the conversation went.
Chapter 11	What's Next?	Individual provided quick preview of Respond Phase.
	Mid-Program Surveys	Individual completes mid-treatment assessment

Coach Meeting 3: Understand Call

Respond Phase

Chapter 12	Power of Acceptance	Individuals learn about how acceptance and change may improve core issue.
	Making Changes	Individuals learn about importance of self-change.
Chapter 13	Changing your Patterns	Individuals are provided tips on how to change the couple's pattern of communication. They then select up to three things they and their partner can change in their patterns during and after a fight.
	Ideas to Improve your Core Issue	Individuals identify changes they and their partner can make to better handle the core issue itself.
	Changing your Core Issue	Individuals write down the changes they would like to make in their own behavior or way of handling the core issue. They also write down changes they would like their partner to make in how he/she handles the core issue.
Chapter 14	Changing your Partner's Core Issue	If the couple has 2 core issues, individuals identify changes they and their partner can make in their patterns of communication during and after a fight and in the way each partner handles their partner's core issue.
Chapter 15	Respond Conversation- 1 Core Issue	Couples have third conversation to discuss agreed upon changes from Chapter 12
	Respond Conversation- 2 Core Issues- First	If couples have 2 core issues, they complete their first Respond conversation to discuss agreed upon changes from Chapter 12 for first partner's core issue.
	Respond Conversation- 2 Core Issues- Second	If couples have 2 core issues, they complete their second Respond conversation to discuss agreed upon changes from Chapter 12 for second partner's core issue.

Chapter 16	Acting on your plan	Individual is provided tips on how to avoid common traps that couples often face when attempting to make changes in their relationship.
	Where Do Things Stand?	Individual completes assessment to track changes during the program
	What's Next for Us?	Individuals are provided feedback on changes they have made during the program. Then, they are asked to write about the nicest thing their partner has done for them since starting their program. Individuals have the option to let the program send this directly to their partner. Finally, they identify what other methods (e.g., couple therapy, continue to utilize program) they would like to do next to strengthen their relationship.
	Post-Program Survey	Individuals complete post-treatment survey.

Coach Meeting 4: Respond Call

Table E.3. Detailed summary of implementation evaluation findings

Element	Research question	Measures	Outcome
Intervention Group			
Fidelity	Were all intended workshop components offered and for the expected duration?	<ul style="list-style-type: none"> Percent of time that server was not available for use 	<ul style="list-style-type: none"> The server was available for use 99.99% of the time.
Fidelity	Were there adaptations to the workshops over the course of the study?	<ul style="list-style-type: none"> Server logs indicating new uploaded content for workshops Project tracking logs noting changes made to workshop content 	<ul style="list-style-type: none"> During Phase 1, we released the Spanish program to accommodate Spanish-speaking couples that were more comfortable receiving help in Spanish (n=16). Detailed automated emails were added to Phase 2 There were minor adaptations made to the project in Phase 1 and Phase 2 simply to better facilitate user experience.
Fidelity	Were individual service contacts delivered reliably across providers to those in the intervention group?	<ul style="list-style-type: none"> Project tracking logs indicating when each coach completed each stage of training Project tracking logs documenting any coaches asked to leave the project for quality control purposes Project tracking logs documenting any protocol deviations (self-reported by coaches in their required post-call notes) related to content covered in individual service contacts 	<p>COACH: Initial Training</p> <ul style="list-style-type: none"> Coach Training Phase 1: Four coaches were trained in August 2016 and three coaches were trained in July 2017. Coach Training Phase 2: Five coaches were trained in December 2017, five coaches were trained in December 2018, four coaches were trained in June of 2019, and four coaches were trained in August of 2019. <p>COACH: Ongoing Supervision</p> <p>Coaches submitted tapes every week for individual supervision and attended weekly group supervision.</p> <p>COACHES ASKED TO LEAVE</p> <ul style="list-style-type: none"> In total, three coaches have been asked to leave the program because they were not able to deliver the interventions with fidelity. Two were asked to leave in January 2019 and one was asked to leave in January of 2020. <p>PROTOCOL DEVIATIONS</p> <ul style="list-style-type: none"> Throughout Phases 1 and 2, only one deviation during individual service contacts with couples in intervention groups was documented: in Phase 1, a couple assigned to OR was mistakenly given an ePREP call.

Element	Research question	Measures	Outcome
Dosage	What percentage of couples who enrolled in the intervention completed 100% of the intervention (i.e., workshops, individual service contacts)?	<ul style="list-style-type: none"> Percentage of couples completing all workshops Percentage of enrolled participants attending all planned individual service contacts <ul style="list-style-type: none"> Note: planned service contacts only occurred for the Coach group 	<p>PERCENTAGE OF COUPLES COMPLETING ALL WORKSHOPS</p> <p>Phase 1: OR: 68% ePREP: 69% Completion rates did not differ significantly across the two conditions using a cutoff of $p < .05$ ($\chi^2(1) = .168$, $p = .682$).</p> <p>Phase 2: OR Coach: 63% OR Contingent Coach: 53% OR Automated Emails: 53% ePREP Coach: 69% ePREP Contingent Coach: 72% ePREP Automated Emails: 63% Completion rates differed significantly across conditions ($\chi^2(5) = 46.39$, $p = .000$). Specifically, completion rates for ePREP Coach were significantly higher than OR Coach ($\chi^2(1) = 4.54$, $p = .033$). Within OR, completion rates for OR Coach were significantly higher than OR Contingent Coach ($\chi^2(1) = 7.72$, $p = .005$) and OR No Coach ($\chi^2(1) = 7.83$, $p = .005$) and OR Contingent Coach and OR No Coach did not differ from each other ($\chi^2(1) = .00$, $p = .993$). Within ePREP, completion rates did not differ significantly across the three conditions using a cutoff of $p < .05$ ($\chi^2(2) = 5.48$, $p = .064$).</p> <p>PERCENTAGE OF PARTICIPANTS ATTENDING ALL PLANNED INDIVIDUAL SERVICE CONTACT</p> <p>Phase 1: OR Coach: 82% ePREP Coach: 72%</p> <p>Phase 2: OR Coach: 69% ePREP Coach: 66%</p>

Element	Research question	Measures	Outcome
Dosage	What percentage of couples who enrolled in the intervention completed 2/3 of the intervention (i.e., workshops, individual service contacts)?	<ul style="list-style-type: none"> Percentage of couples completing at least 2/3 of workshops Percentage of enrolled participants attending at least 2/3 of planned individual service contacts Note: planned service contacts only occurred for the Coach group 	<p>PERCENTAGE OF COUPLES COMPLETING AT LEAST 2/3 OF WORKSHOPS</p> <p>Phase 1: OR Coach: 73% ePREP Coach: 71%</p> <p>Phase 2: OR Coach: 67% OR Contingent Coach: 58% OR Automated Emails: 64% ePREP Coach: 75% ePREP Contingent Coach: 75% ePREP Automated Emails: 68%</p> <p>PERCENTAGE OF PARTICIPANTS COMPLETING AT LEAST 2/3 OF PLANNED INDIVIDUAL SERVICE CONTACTS</p> <p>Phase 1: OR Coach: 82% ePREP Coach: 72%</p> <p>Phase 2: OR Coach: 67% ePREP Coach: 67%</p>
Quality	What was the quality of staff—participant interactions during the individual service contacts?	<p>Average alliance rating, calculated as the average of three alliance questions completed by participants after each coach call: “Overall, this call with my coach was helpful”; “On this call, I felt heard, understood, and respected by my coach”; and “On this call, we talked about what I wanted to talk about”. Response options are 1) Extremely true; 2) Very true; 3) True; 4) Somewhat true; 5) Not at all true; 6) I didn’t have a coach or coach call.</p> <ul style="list-style-type: none"> Note: These questions were only asked for Phase I 	<p>Phase 1: OR Coach: $M = 3.94$ $SD = .93$ ePREP Coach: $M = 4.18$ $SD = .86$</p> <p>Items were reverse scored before creating average alliance score and alliance was significantly higher for ePREP than OR ($t(245) = 2.04$, $p = .042$).</p>

Element	Research question	Measures	Outcome
Engagement	What was intervention couples' overall feedback on the program?	Participant ratings on two items on the post-program survey: "Overall, how much would you say that this program has helped you?" (Response options are 1) A lot; 2) Some; 3) Not at all) and "If a friend were in need of similar help, would you recommend our program to him or her?" (Response options are 1) No, definitely not; 2) No, I don't think so; 3) Yes, I think so; 4) Yes, definitely).	<p>Q1: "Overall, how much would you say that this program has helped you? (1 = a lot, 3 = not at all)"</p> <p>Phase 1: OR: $M = 1.44$ $SD = .50$ ePREP: $M = 1.51$ $SD = .50$ Ratings did not differ across the two programs ($t(442) = -1.40, p = .162$).</p> <p>Phase 2: OR Coach: $M = 1.50$ $SD = .53$ OR Contingent Coach: $M = 1.53$ $SD = .51$ OR Automated Emails: $M = 1.49$ $SD = .52$ ePREP Coach: $M = 1.50$ $SD = .48$ ePREP Contingent Coach: $M = 1.48$ $SD = .44$ ePREP Automated Emails: $M = 1.54$ $SD = .51$ Ratings did not differ across the six programs using a cutoff of $p < .05$ ($F(5,918) = .28, p = .923$).</p> <p>Q2: "If a friend were in need of similar help, would you recommend our program to him or her? (1 = no, definitely not, 4 = yes, definitely)"</p> <p>Phase 1: OR: $M = 3.56$ $SD = .51$ ePREP: $M = 3.56$ $SD = .45$ Ratings did not differ across the two programs using a cutoff of $p < .05$ ($t(442) = -.07, p = .943$).</p> <p>Phase 2: OR Coach: $M = 3.55$ $SD = .53$ OR Contingent Coach: $M = 3.49$ $SD = .49$ OR Automated Emails: $M = 3.48$ $SD = .54$ ePREP Coach: $M = 3.58$ $SD = .46$ ePREP Contingent Coach: $M = 3.63$ $SD = .45$ ePREP Automated Emails: $M = 3.53$ $SD = .51$ Ratings did not differ across the six programs using a cutoff of $p < .05$ ($F(5,918) = 1.64, p = .146$).</p>

Element	Research question	Measures	Outcome
Context	What outside services were utilized by couples in the intervention group during the evaluation period?	<ul style="list-style-type: none"> Percentage of couples who used outside services during the evaluation period For <u>Phase I</u>, participant responds to this question at the end of the intervention: "Which of the following resources have you used since starting the research study?" Response options are: 1) A self-help book or DVD; 2) A different relationship website (not our site); 3) An in-person class or retreat; 4) Individual therapy; 5) Couples therapy. For <u>Phase II</u>, the question was revised to "In the past two months, have you and your partner gone to couple therapy or any type of relationship education program OTHER than the OurRelationship/PREP Online programs?" Response options are (1) Yes; 2) No. 	<p>Phase 1:</p> <p>21.8% of the OR couples and 21.7% of the ePREP couples reported having used at least one of the following resources during the intervention period.</p> <p>1) A self-help book or DVD OR Coach: 10.7% ePREP Coach: 8.7%</p> <p>2) A different relationship website (not our site) OR Coach: 1.4% ePREP Coach: 2%</p> <p>3) An in-person class or retreat OR Coach: .4% ePREP: .6%</p> <p>4) Individual therapy OR Coach: 4.4% ePREP Coach: 4.7%</p> <p>5) Couples therapy OR Coach: 6.7% ePREP Coach: 7.3%</p> <p>Phase 2:</p> <p>The percentage of couples reported having gone to couple therapy or any type of relationship education program other than the OurRelationship/PREP Online programs during the intervention period were 2% in OR Coach, .9% for ePREP Coach, 1.4% in OR contingent Coach, 1.4% in OR No Coach, 1.0% in ePREP contingent Coach, and 1.4% in ePREP No Coach.</p>

Element	Research question	Measures	Outcome
Control / Comparison Group			
Fidelity	Was the initial Welcome call offered to all couples in the control/comparison group and for the expected duration?	<ul style="list-style-type: none"> Percentage of couples who attended single planned individual service contact Average individual service contact duration, calculated as the average of the session lengths (in minutes) 	<p>Attended one individual service contact</p> <p>Phase 1: 100%</p> <p>Phase 2: 100%</p> <p>Average duration of individual service contact</p> <p>Phase 1: 9.35 Minutes.</p> <p>Phase 2: 10.19 minutes</p>
Fidelity	Were individual service contacts delivered reliably across providers to those in the control/comparison group?	<ul style="list-style-type: none"> Project tracking logs indicating when each coach completed each stage of training List of coaches asked to leave the project for quality control purposes as documented in the project tracking logs List of protocol deviations as documented in the project tracking logs (self-reported by coaches in their required post-call notes) related to content covered in individual service contacts 	<p>COACH: Initial Training</p> <p>Coach Training Phase 1: Four coaches were trained in August 2016 and three coaches were trained in July 2017.</p> <p>Coach Training Phase 2: Five coaches were trained in December 2017, five coaches were trained in December 2018, four coaches were trained in June of 2019, and four coaches were trained in August of 2019.</p> <p>COACH: Ongoing Supervision</p> <p>Coaches submitted tapes every week for individual supervision and attended weekly group supervision.</p> <p>COACHES ASKED TO LEAVE</p> <p>In total, three coaches have been asked to leave the program because they were not able to deliver the interventions with fidelity. Two were asked to leave in January 2019 and one was asked to leave in January of 2020.</p> <p>PROTOCOL DEVIATIONS</p> <p>Throughout Phases 1 and 2, no deviation during individual service contacts with couples in the control group was reported.</p>

Element	Research question	Measures	Outcome
Fidelity	What outside services were utilized by couples in the control group during the evaluation period?	<ul style="list-style-type: none"> Participant responds to this question at post, 4-month follow-up, and 6-month follow-up: "Which of the following resources have you used since starting the research study?" Response options are: 1) A self-help book or DVD; 2) A different relationship website (not our site); 3) An in-person class or retreat; 4) Individual therapy; 5) Couples therapy. Note: This question was not asked in Phase II. 	<p>Phase 1:</p> <p>The percentage of control couples who reported having used at least one of the following resources were 13.4%, 14.6% and 16.4% at post, 4, and 6 months after randomization respectively.</p> <p>1) A self-help book or DVD: post: 5.7%; 4 months: 5.7%; 6 months: 8.1%</p> <p>2) A different relationship website (not our site) post: 1.8%; 4 months: 2.4%; 6 months: 1.4%</p> <p>3) An in-person class or retreat post: .6%; 4 months: .6%; 6 months: .8%</p> <p>4) Individual therapy post: 3.8%; 4 months: 5.3%; 6 months: 5.7%</p> <p>5) Couples therapy. Post: 2.2%; 4 months: 2.0%; 6 months: 2.2%</p>