Impact Evaluation of Strengthening Relationships/ Strengthening Families (SR/SF) in Central Texas

Final Impact Evaluation Report for

Texas State University

November 13, 2020

Prepared by

Andrea Hutson, PhD, Agile Analytics

Norma J. Perez-Brena, PhD, Texas State University

Michelle Toews, PhD, Kansas State University

Jeneé Duncan, PhD, Texas State University

Renee Perez, LPC, Texas State University

Recommended Citation:

Hutson, A., Perez-Brena, N. J., Toews, M., Duncan, J., & Perez, R. (2020). Impact evaluation of Strengthening Relationships/Strengthening Families (SR/SF) in Central Texas. Final Impact Evaluation Report.

Acknowledgements:

We are grateful to the families and youth who participated in this project, and to the following school districts and non-profit organizations who collaborated and served as part of our advisory board: Austin Independent School District, Comal Independent School District, Del Valle Independent School District, Hays Consolidated Independent School District, San Marcos Consolidated Independent School District, SAFE Alliance, and Hays-Caldwell Women's Center. We thank Kristin Anders, Shelby Astle, Elia Bueno, Allison Gilman Caplovitz, Sarah Colburn, Alicia Collins, Carolina deJongh, Gabriel Dhesse, Eric Goodcase, Lara Hoss, Kimberlyn Jaggers, Heather Kristoff, Lindsay Lamb, Julie Moore, Elizabeth Norton, Paloma Ocampo, Veronica Panduro, Roxana Perez, Eddie Reyes, Denise Rovzar, Lindsay Rudy, Haleigh Williams, and the countless undergraduate interns and research assistants for their assistance in conducting this investigation. We would also like to thank our collaborator, Mark Fienberg, our program officer, Meghan Heffron, and evaluation technical assistant, Angela D'Angelo for their feedback throughout this project.

Disclosure:

We do not have any conflicts of interest, financial or otherwise.

This publication was prepared under Grant Number 90FM0067-02-00 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: "Impact Evaluation of Strengthening Relationships/Strengthening Families (SR/SF) in Central Texas"

Objective. The purpose of this study was to assess the impact of a relationship-based intervention program serving adolescent parents in Central Texas. The primary goal was to assess the effects of receiving relationship and co-parenting curricula (the 'full-treatment') compared to only receiving either a relationship or co-parenting curriculum plus a control curriculum (the 'partial treatment') on participants' coparenting and relationship attitudes and behaviors. Our secondary goals were (1) to examine the long-term impacts on the above outcomes and on participants' adjustment, and (2) to examine the impacts of each of the curricula individually on the same outcomes at mid-test.

Study Design. A quasi-experimental design was used such that eight schools were block randomized into different treatment conditions. Individuals who were identified as adolescent parents were invited to participate in the program, although couples could join together. Participants' outcomes were assessed at pre-test, mid-test (after receiving curriculum 1), post-test (after receiving curriculum 2), and three months after post-test.

Results. When answering the primary research questions, we found there were no statistically significant differences between the full and partial treatment groups, with one exception. Students in the full treatment group reported less coparenting conflict than their counterparts. Regarding our exploratory research questions, no group differences emerged at the three-month follow up. When comparing results for the individual curriculum compared to the control curriculum, no statistical differences emerged at mid-test. However, one outcome approached significance; specifically, students who attended the coparenting curriculum reported slightly more communication with the other biological parent of their child.

Conclusion. Our inability to find significant differences between groups may have been due to key implementation challenges, such as students' rate of completion of the two-semester model and the need to adapt lessons to address students' attendance challenges (e.g., online lessons, make-up lessons). Additionally, concepts taught in both curricula overlapped; thus, gains from having two semesters of content may have been marginal. Finally, students began the study with healthy self-reported attitudes and behaviors potentially leading to ceiling and floor effects. Lessons learned include: the need to find alternate delivery methods to address barriers to participation, the need to reduce program lengths, the value of incentives in program completion, and the need to collaborate with multiple school types.

Contents

Tables

Figures

IMPACT Evaluation of Strengthening relationships/Strengthening families (SR/SF) in Central Texas

I. INTRODUCTION

A. Introduction and study overview

The United States has the highest teen pregnancy rate in the industrialized world (Sedgh et al., 2015). In fact, in 2017, 194,377 adolescents under the age of 20 became parents (Martin et al., 2018). Texas has the fourth highest teen pregnancy and birth rates in the country (Center for Disease Control and Prevention [CDC], 2018) and in Central Texas alone, 1,711 adolescents under the age of 18 became pregnant and gave birth in 2015 (Texas Department of State Health Services (TDSHS), 2015). Of the more than 1,700 adolescents who gave birth in Central Texas, 95% identified as Latino (TDSHS, 2015).

Adolescent pregnancy and parenting, particularly for low-income, ethnic minorities, often result in adverse health, psychosocial, and socioeconomic consequences for the mother, father, and children such as greater health risks, poorer psychological functioning, decreased rates of school completion, higher levels of relationship instability, increased risk of unemployment, greater dependence on welfare, and so forth (see Hoffman & Maynard, 2008). Therefore, it is critical to implement programs that provide these at-risk youth with services that enhance their well-being and strengthen their families. In response to this need, a plethora of programs that emphasize self-sufficiency skills have been implemented. For example, in Central Texas, Pregnancy, Education, and Parenting (PEP) programs were implemented over 20 years ago to enable adolescent parents to become self-sufficient, responsible, job-oriented citizens. Although programs such as PEP provide numerous services, including career counseling and job-readiness training, they do not offer relationship education, a major factor in enhancing well-being and strengthening families (Waite & Gallagher, 2000). The directors of local PEP programs report this as a major limitation of their program. The adolescent parents enrolled in PEP also express a strong interest in learning how to improve their relationships.

To address this need, the purpose of our project was to adapt, implement, and evaluate our Strengthening Relationships/Strengthening Families (SR/SF) program designed to provide pregnant and parenting adolescents with healthy relationship and co-parenting skills in order to improve their well-being and family functioning. Specifically, we extended our previous program, which used Love Notes, a skill-based relationship education program designed specifically for young parents, to include a modified version of Family Foundations, a skill-based co-parenting curriculum. We chose to incorporate coparenting into our existing program (See Appendix A for our Logic Model) because, although our previously funded program was successful in improving adolescent parents' communication and conflict resolution skills,

increasing adolescents' ability to set goals for themselves and their relationships, and reducing intimate partner violence (Toews & Yazedjian, 2010; Toews et al., 2011), our participants continued to need more assistance negotiating the coparent relationship. Further, extant research suggests the co-parenting relationship is more directly linked to reduced intimate partner violence and maternal depression, while more closely linked to increased paternal involvement. These findings are in line with previous research regarding the related, but unique, needs and impacts of romantic and co-parenting relationships on family functioning.

The goal of our project was to provide adolescent parents with a comprehensive relationship education program that strengthens the multiple relationships they must manage as new parents to ultimately enhance their well-being and strengthen their families. To reach this goal, four facilitators implemented the program during the school day with the assistance of interns. The program was delivered to adolescents enrolled in high school and receiving Pregnancy, Education, and Parenting (PEP) or social support services in Central Texas. We conducted concurrent sessions at seven-to-nine high schools each semester. Groups of 10-20 pregnant and parenting adolescents met at each school every week to participate in sessions focusing on healthy relationships, co-parenting, or life-skills (our control group for impact evaluation). These sessions covered topics such as healthy and unhealthy relationships, communication skills, conflict resolution strategies, negotiating childrearing strategies, and managing and sharing parenting duties. We used this approach because it incorporated promising practices for relationship education programs serving low-income, ethnically diverse populations. Specifically, our program was long-term, provided opportunities to practice interpersonal skills, employed interactive teaching methods, reduced barriers to participation, partnered with organizations that were well-established in the community, and incorporated materials that were age- and culturally-appropriate.

Students enrolled in the evaluation received two semesters of services – full treatment (relationship and co-parenting curricula) or partial treatment (relationship or co-parenting curricula and the control curricula). Outcomes were measured between semesters (the mid-test), after two semesters (post-test) and three months later (long-term).

The primary goal of this impact study was to examine the effect of experiencing both the relationship and co-parenting curricula (the 'full-treatment' condition) compared to only receiving one of the curricula plus a control curriculum (the 'partial treatment' condition) on pregnant and parenting adolescents' attitudes about coparenting and relationships, their communication skills, and their relationship behaviors at the end of treatment (post-test). Our secondary exploratory goals were (1) to examine the long-term impacts on the above outcomes, (2) measure long-term changes in students' levels of anxiety and depression 3-months post-treatment, and (3) examine the impacts of each of the curricula on the relationship and coparenting attitudes and behaviors at mid-test.

In parts B and C of this section, we present the primary and secondary research questions. In Section II, we introduce the intervention and counterfactual conditions, and provide information about how the intervention was structured and conducted in the study. In Section III, we explain

more about how the study was designed and how data were collected. In Section IV, we discuss our analysis strategy and provide information about the outcome measures used in the study. In Section V, we present the results of the study, and in Section VI, we provide discussion surrounding those results. Additional information about the study is provided in the references and the appendices.

B. Primary research questions

RQ1 Do pregnant and parenting adolescents who participate in the full-treatment group (i.e., receiving relationship and co-parenting curricula) report healthier co-parenting attitudes compared to those adolescents who participate in the partial-treatment group (i.e., receiving one of the curricula plus a control curriculum) at post-test?

RQ2 Do pregnant and parenting adolescents who participate in the full-treatment group report healthier co-parenting communication compared to the partial-treatment group at post-test?

RQ3 Do pregnant and parenting adolescents who participate in the full-treatment group report healthier relationship attitudes compared to the partial-treatment group at post-test?

RQ4 Do pregnant and parenting adolescents who participate in the full-treatment group report healthier relationship behaviors compared to the partial-treatment group at post-test?

C. Secondary research questions

Three-month follow up questions

RQ5 Do pregnant and parenting adolescents who participate in the full-treatment group report healthier co-parenting attitudes compared to those adolescents who participate in the partial-treatment group three-months post-treatment?

RQ6 Do pregnant and parenting adolescents who participate in the full-treatment group report healthier co-parenting communication compared to the partial-treatment group three-months post-treatment?

RQ7 Do pregnant and parenting adolescents who participate in the full-treatment group report healthier relationship attitudes compared to the partial-treatment group three-months post-treatment?

RQ8 pregnant and parenting adolescents who participate in the full-treatment group report healthier relationship behaviors compared to the partial-treatment group three-months post-treatment?

RQ9 Do pregnant and parenting adolescents who participate in the full-treatment group report reduced RQ depressive symptoms compared to the partial-treatment group three-months post-treatment?

RQ10 Do pregnant and parenting adolescents who participate in the full-treatment group report reduced worry compared to the partial-treatment group three-months post-treatment?

Mid-test questions

RQ11 At the end of the first semester, do pregnant and parenting adolescents who received our coparenting curriculum report healthier co-parenting attitudes compared to those who did not receive either of the treatment curricula at mid-test?

RQ12 At the end of the first semester, do pregnant and parenting adolescents who received our coparenting curriculum report healthier co-parenting communication compared to those who did not receive either of the treatment curricula at mid-test?

RQ13 At the end of the first semester, do pregnant and parenting adolescents who received our healthy relationships curriculum report healthier relationship attitudes compared to those who did not receive either of the treatment curricula at mid-test?

RQ14 At the end of the first semester, do pregnant and parenting adolescents who received our healthy relationships curriculum report healthier relationship behaviors compared to those who did not receive either of the treatment curricula at mid-test?

II. INTERVENTION AND COUNTERFACTUAL CONDITIONS

This section provides a description of the intervention being evaluated and the services that are intended as the comparison to the intervention, including any "business as usual" resources.

A. Description of program as intended

Intended components: This is a multi-component intervention in which pregnant and parenting adolescents receive two of three curricula: Love Notes (Healthy Relationships/HRR), From Teen Parent to Team Parent (Coparenting/COPAR), and Life After Graduation (Life & Financial Skills/CONT). A block randomization design was used to randomly assign the eight schools into one of four groups. All groups received the co-parenting (COPAR) and healthy relationships (HRR) curricula, along with the Life After Graduation control (CONT) curriculum across the five school years. The order in which a school received the curricula varied by group membership and by year, with only two curricula delivered during one school year. Students received two semesters of service as part of the program design, and their treatment assignment was based on the school and year in which they enrolled.

The Love Notes (Healthy Relationship/HRR) curriculum is a research-based intervention focused on teaching youth healthy relationships skills such as communication skills, conflict management, knowledge about the benefits of marriage, stress and anger management, affection and intimacy, problem solving, and negotiation skills. This curriculum was created for adolescents and has been shown to effectively improve participants' knowledge, attitudes, and behaviors associated with healthy relationships (Barbee et al., 2016; Kerpelman, 2010).

The From Teen Parent to Team Parent (Coparenting/COPAR) curriculum is an adapted curriculum based on the Family Foundations curriculum. Family Foundations is an evidence-based program aimed at improving coparenting by teaching relationship skills such as healthy communication, conflict and anger management, intimacy and affection, parenting goals, coparenting problem solving, and managing multiple coparents. This curriculum has been shown to increase coparental support, parental efficacy, and positive child adjustment, and reduce harsh parenting and parental stress (Feinberg & Kan, 2008; Feinberg et al., 2010; Feinberg et al, 2016). The curriculum has been adapted to serve minority (primarily African American) adolescent parents (Lewin, et al., 2012) and SR/SF adapted it further for Latinx youth in a school-based setting. In collaboration with the curriculum developer, focus groups were held in December 2015 with adolescent parents, their own parents, and school staff to identify coparenting challenges and strengths. These focus groups informed initial adaptations which were piloted in Spring 2016. Participants provided feedback on each adapted lesson using a short survey, and they provided general feedback on the curriculum during focus groups at the end of the Spring 2016 semester. The curriculum developer, SR/SF program directors, SR/SF program coordinators, and the local evaluator used this feedback to make final edits to the curriculum before the impact evaluation began.

The Life After Graduation (CONT) curriculum was selected by the program to serve as an "attentional control" which allowed participants to receive the same amount of positive

interaction from the program facilitators and peers without receiving the actual treatment. This high-quality control curriculum was focused on education, career, and financial literacy topics. By hosting an attentional control group, SR/SF was able to host sessions during the assigned time, encouraging students to participate and preventing another program from using that same time to teach participants about potentially overlapping topics. We did not employ a "usual or no services" control group in this project.

Intended content: See Table 1 for the intended content of each curriculum.

Table 1. Key features of each curriculum

Love Notes (Healthy Relationship Education, HRR)	From Teen Parenting to Team Parenting (Coparenting Education, COPAR)	Life After Graduation (Life Skills Education- control. CONT)
1. Knowing Yourself	**1.** Introduction to team parenting	**Financial Literacy & Life Skills**
2. My Expectations—My Future	**2.** Taking Care of Yourself and your Relationships	**1.** Financial Literacy
3. Attractions and Starting Relationships	**3.** Teamwork and listening skills	**2.** Budgeting – Cost of Parenting
4. Principles of Smart Relationships	**4.** Conflict, your body, and your child	**3.** Smart Shopping
5. Is it a Healthy Relationship	**5.** Managing conflict and thoughts	**4.** Cars
6. Dangerous Love	**6.** Working it out	**5.** Housing
7. Decide, Don't Slide!	**7.** Security and problem solving	**Career Pathways After Graduation**
8. Let's Talk About Sex	**8.** Discipline and problem solving methods	**6.** Workforce
9. What's Communication Got to Do With It	**9.** Other Partners	**7.** Education After High School
10. Communication Challenges and More Skills	**10.** Grandparents	**Life Balance & Planning**
		8. Time Management
		9. Nutrition & Health
		10. Birth Control & Family Planning

Full treatment condition: The full treatment condition was comprised of students who were assigned to both the coparenting (COPAR) and Healthy Relationship (HRR) curricula. The partial treatment condition was comprised of students who were assigned to either the COPAR or HRR curriculum and the control (CONT) curriculum. See Table 2 for a description of the curriculum schedule per treatment group. More information will be presented in Section III.1.

Table 2. Intended component per treatment condition assignment

School Year & Semester		Randomized Treatment Groups (two schools per group)			
		G1	G2	G3	G4
16/17	FALL	COPAR	CONT	CONT	HRR
	SPR	HRR	HRR	COPAR	COPAR
17/18	FALL	CONT	COPAR	HRR	CONT
	SPR	COPAR	CONT	CONT	HRR
18/19	FALL	HRR	HRR	COPAR	COPAR
	SPR	CONT	COPAR	HRR	CONT
19/20	FALL	COPAR	CONT	CONT	HRR
	SPR **[COVID-19 Pandemic]**	HRR	HRR	COPAR	COPAR
2020	FALL	Final Report Due			

Note. Dark blocks indicate the full treatment group, Light grey blocks indicate the partial treatment (Healthy relationships only) groups, and white boxes indicate the partial treatment (Co-parenting only) group. Three-month activities are for student retention purposes. Students participate in the study for two consecutive semesters and receive the curricula that were randomly assigned by school.

Planned dosage and implementation schedule: Each curriculum was comprised of ten one-hour lessons that were delivered weekly across one school semester (about 4.5 months). Students were assigned to two curricula; thus, their total participation time was one school year, or approximately 9 months.

Intended delivery: After schools students were assigned into one of the four groups, Project Facilitators (PFs) were assigned to work with one group (two schools) for the duration of the grant. The program has had seven PFs in total, and one of the PFs has changed schools, such that all groups have had one change in their PF over the course of the grant. We also left and replaced two schools. Additionally, an intern was assigned to every campus each school year. Interns provided support for PFs, helped deliver lessons, and provided support to students while they learned key skills.

Lessons were generally delivered in the classroom by the PF with support from the intern assigned to the campus. Although it was not part of the original study design, virtual methods of attending lessons were added to the program in 2018 and were used throughout the rest of the study. These methods were added to support students who were ill (or whose children were ill), and those who were not able to come to school for other reasons. These methods are explored in Section II.D, below. Students who were not able to attend the lesson either in person or virtually received a one-on-one make up lesson from the PF.

Target population: Participants were required to be (1) an adolescent/young adult between the ages of 14-20; (2) enrolled in one of our target schools; (3) be pregnant, parenting, or have a partner who was pregnant (for male participants). Some of the students in the study were in romantic relationships with each other, and thus participated as a couple. However, the program was delivered individually, and being in a couple was not required. Indeed, the vast majority of students (>75% overall) participated individually.

Education and training of staff: The Project Directors provided oversight of the entire intervention program, as well as the development and implementation of all policies, procedures, and processes. The Project Coordinators (PCs) trained and provided ongoing support to PFs and interns. PFs and interns were given extensive training on each curriculum at the beginning of each semester. Most PFs led groups for the SR/SF program for multiple years and delivered each curriculum several times. See Table 3 for more training information.

SR/SF staff minimized the opportunities for treatment cross-over effects by ensuring PFs and interns did not transfer to other treatment groups during a given semester. Further, PCs conducted monthly site visits to observe PFs and interns and to note any discrepancies to program content or delivery. Additionally, program delivery goals and questions about content delivery were discussed at bimonthly team meetings to ensure that unintentional treatment cross-over effects did not emerge in the classroom.

All data collection and implementation staff received training regarding adverse events (primarily with regard to domestic violence or child abuse) during their initial staff training for this project, as well as ongoing supervision around these issues. In addition, evaluation team members were coached to respond to participant embarrassment or discomfort with data collection or intervention in an appropriate and compassionate manner.

Table 3. Staff training and development to support intervention and counterfactual components

Component	Education and initial training of staff	Ongoing training of staff
Healthy Relationships	Facilitators are female and hold at least a bachelor's degree in human development and family science or related majors and received four days of initial training.	Facilitators receive a half-day of semi-annual refresher training in the intervention's curricula from study staff. Facilitators receive an additional 2-3 days of supplemental training in key topics related to implementation (e.g., workshop delivery, online delivery, supervision)
Coparenting	Facilitators are female and hold at least a bachelor's degree in human development and family science or related majors and received four days of initial training.	Facilitators receive a half-day of semi-annual refresher training in the intervention's curricula from study staff. Facilitators receive an additional 2-3 days of supplemental training in key topics related to implementation (e.g., workshop delivery, online delivery, supervision)
Life After Graduation	Facilitators are female and hold at least a bachelor's degree in human development and family science or related majors and received four days of initial training.	Facilitators receive a half-day of semi-annual refresher training in the intervention's curricula from study staff. Facilitators receive an additional 2-3 days of supplemental training in key topics related to implementation (e.g., workshop delivery, online delivery, supervision)

B. Description of counterfactual condition

The counterfactual condition was comprised of students who received the control curriculum (CONT) plus the healthy relationships curriculum (Partial-HRR), or the control curriculum plus the co-parenting curriculum (Partial-COPAR).

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

We also developed implementation-based research questions to understand if our program was implemented as intended, and the potential contextual factors that might have prevented us from implementing as intended in some or all sites (Table 4). This information was used to inform our impact evaluation, as it could provide context for understanding why our program succeeded OR what could have led to us seeing little to no impact in program effects. Of course, we acknowledge that true causation cannot be assumed from this data.

Table 4. Implementation Research Questions

Implementation element	Research question
Fidelity	• Were all intended intervention components offered, per curriculum? • What were the unplanned adaptations to key intervention components, per curriculum?
Dosage	• How many students attended the two-curricula model, per intervention assignment? • How many classes did students attend, on average, per curriculum?
Quality	• What was the quality of staff–participant interactions, per curriculum?
Engagement	• How engaged were intervention group members in the intervention, per curriculum?
Context	• What external events affected implementation?

D. Services actually received by the intervention and control/comparison groups

Sections A-C offer a description of our intended design. In general, we administered our program (e.g., intended curricula, lessons #, curricula schedule) as intended with three exceptions. We describe these three exceptions early in this report because they will inform various methodological decisions noted across the report, such as a change in the study design from a randomized-controlled trial (RCT) to a quasi-experimental design (QED).

Replacement of Schools. In 2018, we left one of our original schools due to low enrollment, and replaced this school with a new school that showed an increase in pregnant and parenting adolescents. In 2019, we left a second school and replaced it with a new school for similar reasons. These school changes compromised our original RCT; thus, requiring we analyze the study as a QED.

Virtual Lessons. In order to ensure program participation and retention, we offered lessons virtually to absent students. This virtual and synchronous education model was not part of our original design. There were two primary methods to reach students:

- The "Homebound Heroes" program was introduced in the Fall of 2018. In this method, a student in the classroom, or the program intern, contacted individual students at home via cell phone. The student was then able to listen to the presentation as it happened, participate in the discussion, and complete the activities at home.
- The virtual program was introduced around the same time. In this method, a student participated in lessons via an online meeting platform (e.g., Zoom). PFs set up a webcam and microphone in the classroom so that students could see the presentation and activities. Students could thus participate to a large extent at home.

SR/SF staff made a concerted effort to perfect these virtual methods of student lesson delivery; in the fall of 2019, evaluation staff conducted focus groups at all schools to determine the benefits and drawbacks to various approaches. This extensive prework proved valuable in the spring of 2020, when the final four SR/SF lessons were delivered virtually to all students due to schools closing because of COVID-19. SR/SF staff were able to immediately pivot and conduct lessons with students and had success in reaching the majority of participants.

Incentive Model. *SR/SF's* goal was to have students complete all ten lessons if possible, and an incentive model was piloted in the spring of 2018 to encourage students to do so. In the incentive model, students received a $50 gift card if they completed 10/10 lessons in a semester without having any make-up lessons. If students missed a lesson, they were eligible for a $10 gift card and a drawing for a $40 gift card if they completed all 10 lessons with make ups. The model worked so well at the pilot campuses that it was then rolled out to all schools in the fall of 2018 and was used for the last two years of the program.

III. STUDY DESIGN

In this section, we describe the study eligibility criteria, research group formation, research design, and data collection methods.

A. Sample formation and research design

This section provides a brief description of the study design and the process for creating intervention and comparison groups.

1. Sample formation

All pregnant or parenting students attending one of our eight schools beginning Fall 2016 were invited to participate in our program and evaluation. Trusted staff at each school (e.g., nurse, social worker, childcare center director) identified these students and provided the initial introduction to our program. Participants were required to be pregnant or parenting adolescent (ages 13-20) mothers and fathers. The intervention was primarily targeted at individuals, although couples were welcomed. Only 22% of the analytic sample for the primary research questions were couples.

Randomization. A block randomization design was used to randomly assign the eight schools into one of four possible treatment sequences. As noted in Table 2, four groups were created to reflect multiple counterbalanced conditions in which each group received the curricula in a different order. Each school was randomly assigned to one of the four treatment sequences at the beginning of the study by project directors who picked names out of a hat. PFs were assigned to one of the four groups and served both schools within that treatment sequence. Students participated in the study for two consecutive semesters and received the curricula that were randomly assigned to their school. That is, individual participants received two of the three curricula over the two semesters. Over time, each school administered all three curricula as they rotated through the sequence, and ultimately there were two possible conditions analyzed:

- **Full treatment:** Received both the coparenting and healthy relationships curriculum
- **Partial treatment**: Received either the coparenting or the healthy relationships curriculum and the control curriculum

In general, pre-test data were collected at the beginning of the first semester, mid-test data were collected at the end of the first semester, post-test data were collected at the end of the second semester, and three-month follow-up data were collected three months after post-test. To assess our research questions, we compared outcomes for the two groups of interest: the full and partial treatment conditions.

Group Changes. In Year 3 (2018-2019) of our study, we moved out of one of our target schools and into a new school due to low student enrollment. Similarly, in Year 4 (2019-2020), we left one school and replaced it with a new school for similar reasons. The new schools were placed in the same treatment group as the school they replaced. These moves required that we treat our

study design as a QED, which required developing a plan for ensuring that the two conditions were well matched at pre-test by establishing baseline equivalence on the groups based on a pre-determined set of variables.

Finally, late in the spring semester of Year 4, the COVID-19 pandemic forced schools to close (mid-March, 2020). All programming was moved online, and the final four lessons were administered virtually to all students. Because of the changes to the program and the significant changes in other parts of students' lives, we suspected that students from this cohort might be systematically different from students in other cohorts. Preliminary analyses confirmed significant differences in coparenting communication, depression, and anxiety at post-test.

Additionally, students affected by COVID were not evenly distributed across all conditions. As can be seen in Table 2, a larger proportion of students were in the full treatment condition during the affected year (19/20) than the two partial treatment conditions. When we added this as a variable to our primary analyses, it was over the threshold for baseline equivalence. For these reasons, student data from this cohort were not analyzed as part of our primary research questions. We did, however, analyze data for this cohort for our secondary research questions taken at mid-test (pre-COVID).

Research Question comparison groups. Below, we list all of the analysis groups for the research questions. See Table 2, above, to see the blocks referenced in this section.

RQ 1-4: We compare post-test results for full treatment (dark grey blocks) to the post-test results for the partial treatment groups (light grey and white blocks)

RQ 5-10: We compare three-month follow-up results for full treatment (dark grey blocks) to the three-month follow-up results for the partial treatment groups (light grey and white blocks)

RQ 11-12: We compare mid-test results for participants who received the coparenting curriculum (COPAR) in their first semester, to those who received the control curriculum (CONT) in their first semester.

RQ 13-14: We compare mid-test results for participants who received the healthy relationship curriculum (HRR) in their first semester, to those who received the control curriculum (CONT) in their first semester.

B. Data collection

In this section we first describe the data collection methods for the implementation analyses, then we discuss the data collection methods for the impact analyses.

1. Implementation analysis

Most of the implementation analyses are descriptive in nature. Therefore, we report summary statistics such as sum scores, percentages, or mean scores. Content analyses were used to review existing program records and tabulate commonly mentioned themes.

Several data sources were used, such as the master calendar, PC observation forms, PF focus group, tracking sheets, nFORM exit survey, participant focus groups, advisory board notes, and HMRF progress reports.

- Master calendar - PF entered their lesson dates and times in an online calendar that was monitored by PCs, Program Directors, and the grant senior secretary.
- PC observations – PCs observed each PF and intern monthly. These observations were meant to help build facilitator skills and monitor program fidelity. Each month, one PC attended a lesson for a single PF. Each PC attends 4 lesson per month to observe each of the 4 PFs and 8 interns across the 8 sites. PCs discussed observations with each other and the program director at a weekly meeting to monitor PF/intern skills, PF/intern growth, and fidelity to the curriculum. Observations occurred 3 times per semester for a total of 6 observations per year.
- PF focus group- at the end of the grant cycle, the local evaluator and Program Director conducted a PF focus group where facilitators were asked to note the # (%) of students who completed online lessons and make-up lessons, per semester. PFs were also asked to recall how many adaptations were made when they facilitated classes (1) online or (2) during make-up lessons. Finally, PFs were asked to recall if adaptations differed for each curriculum. Response ranges and averages were estimated for the quantitative questions, and inductive coding was used to identify common themes in the qualitative questions.
- Tracking sheets – We supplemented nFORM with our own internal attendance record system (called the tracking sheet). Tracking sheets included data on student attendance, if a lesson was delivered in a group setting, through a 1:1 meeting, or virtually. Data in tracking sheets were input by PFs and audited at the end of each semester by PCs. Discrepancies were clarified, and tracking sheets were cleaned after the audit.
- nFORM exits – nFORM – the Information, Family Outcomes Reporting and Management system – is an online client management system mandated by the federal funder. The nFORM exit survey is the federally required survey which is administered at the end of services. This survey assesses participant demographics, relationships, coparenting, marriage, sexual risk behavior attitudes, along with participants' reflection on the program. nFORM exit surveys were collected at the end of two semesters (ideally); however, some students left our program early and those students were asked to complete the exit before they left. Four hundred and two students completed exit surveys (62% completion rate).
- Participant focus groups – All study participants attended a focus group at the end of each semester. The size of each focus group varied by school and semester. However, on average, we collected focus group data from 122 students/semester.
- Advisory board notes – An advisory board meeting was held at the end of each semester. The SR/SF administrative assistant compiled all documents discussed at the meeting and took meeting minutes which served as a written account of each meeting.

- HMRF progress report – Progress reports were created by the Program Directors and submitted quarterly.

Data collection methods and analysis methods are listed in Tables 5 and 6 below.

Table 5. Data used to address implementation research questions

Implementation element	Research question	Data source	Timing/frequency of data collection	Party responsible for data collection
Fidelity	Were all intended intervention components offered?	Master calendar (MC), Program Coordinator (PC) observation form	Daily (MC); monthly (PC observation)	Program facilitators; program coordinator
Fidelity	What were the unplanned adaptations to key intervention components, per curriculum?	Program facilitator focus group	End of grant cycle	Local Evaluator and Director
Dosage	How many students attended the two-intervention model, per intervention assignment?	Tracking sheets	Biannually	Program facilitators
Dosage	How many classes did students attend, on average, per curriculum?	Tracking sheets, master calendar, nFORM hours	Daily; Weekly	Program facilitators
Quality	What was the quality of staff–participant interactions, per curriculum?	PC observation form; nFORM exit survey; participant focus groups	Monthly; at the end of program participation	Program facilitators; Research team; Local Evaluator
Engagement	How engaged were intervention group members in the intervention, per curriculum?	PC observation form; nFORM exit survey-Qualitative Questions	Monthly; at the end of program participation; Biannually	Program facilitators; Research team; Director
Context	What external events affected implementation?	Advisory board notes; HMRF progress report	Biannually	Director

Table 6. Implementation analysis methods

Implementation element	Research question	Measures
Fidelity	Were all intended intervention components offered and for the expected duration?	• Total number of sessions delivered • % of observed lessons that met the 75% lesson content cut off.
Fidelity	What were the unplanned adaptations to key intervention components, per curriculum?	• List of unplanned adaptations due to make ups, and online service delivery. • # students using online service delivery • % of lessons online vs. in-person • # students receiving make-ups

Implementation element	Research question	Measures
Dosage	How many students attended the two-intervention model, per intervention assignment?	• # of students who attended at least one lesson across two consecutive semesters • # (%) of students who completed 6/10 and 10/10 lessons per intervention assignment, across two consecutive semesters
Dosage	How many classes did students attend, on average, per curriculum?	• % of students who complete 6/10 and 10/10 lessons
Quality	What was the quality of staff–participant interactions, per curriculum?	• % of sessions that "met" or "exceeded" expectations in "Lesson Layout" and "Lesson Facilitation" categories. • Students' average response to the utility and satisfaction questions from nFORM (scores range from 1–5) • The # of times a positive comment (vs. negative) about the program was recorded during focus groups
Engagement	How engaged were intervention group members in the intervention, per curriculum?	• % of sessions with high student engagement, calculated as the percentage of observed lessons scored as "meets" or "exceeds" expectations in "Relationship with Participants." • List of qualitative themes that emerge in the nFORM and SR/SF qualitative question regarding students' reaction to the program.
Context	What external events affected implementation?	• # of schools affected by weather events, policy changes, or community issues. • List of weather events, policy changes, or community issues that adversely impacted programming or attendance and length of impact (acute vs. long-term impacts).

2. Impact analysis

Enrollment. Eligible participants were notified of the project via a trusted staff member at the schools (e.g., nurse, social worker, childcare center director). Students were then provided with a recruitment folder that included a recruitment flyer, parental consent form, and adolescent assent forms. In addition, bilingual staff made follow-up calls, sent text messages, or provided in-person reminders to answer any follow-up questions. Students were asked to return the recruitment packet if they chose to participate.

Survey administration. Trained research assistants administered surveys to participants using computer tablets that directed students to a personalized link on Qualtrics, an online survey tool, during a designated class. Students completed a pre-test survey at the start of the semester prior to beginning lessons, and then completed a mid-test and post-test survey after the completion of each semester. Finally, a follow-up survey was completed three months after the completion of the second semester. If students were not in school on the day surveys were administered, or were not enrolled in school (e.g., if a student dropped out of school), then trained research assistants called and invited students to complete their survey online using the same Qualtrics tool. See Table 7 for timing, mode, parties responsible for data collection.

Table 7. Key features of the impact data collection procedures

Condition	Participants	Timing of data collection	Mode of data collection	Party responsible for data collection	Start and end date of data collection (for fall enrollees)
Full Treatment	Intervention group study participants who receive HRR + COPAR	**Enrollment** (pre-test/ beginning of 1st semester) **Mid-test** (4.5 months after pre-test/end of 1st semester) **Post-test** (9 months after pre-test/end of 2nd semester) **3-month follow-up** (12 months after pre-test)	In-person online survey (SR/SF survey) - students take the online survey on tablet devices in class.	Research team	August-September (pre-test) November/December (mid-test) April/May (post-test) August/September (3-month follow-up)
Partial Treatment (Counterfactual / Control)	Comparison participants who receive partial treatment (HRR + CONT or COPAR + CONT)	**Enrollment** (pre-test/beginning of 1st semester) **Mid-test** (4.5 months after pre-test/end of 1st semester) **Post-test** (9 months after pre-test/end of 2nd semester) **3-month follow-up** (12 months after pre-test)	In-person online survey (SR/SF survey)	Research team	August-September November/December April/May August/September

IV. ANALYSIS METHODS

In this section we provide a description of the sample construction, outcome measures, and the baseline equivalence of the treatment (Full) and control (Partial) conditions.

A. Analytic sample

Across the four years of the study, there were seven participating cohorts of students (one for each school semester of the first three years of the study, and one for the first semester of the fourth year of the study. The composition of the schools changed slightly over time (with two new schools entering the study and two schools exiting the study), but there were always eight schools participating. As a whole, there were 56 clusters across the study (7 cohorts of students x 8 schools).

The analytic sample included all intervention students from the first three years of the study (cohorts 1-6) who had both pre- and post-test data (Research Questions 1-4), all intervention students from the first three years of the study (cohorts 1-6) who had both pre-test and 3-month follow-up data (Research Questions 5-10), and all intervention students for all four years of the study (cohorts 1-4) who had both pre-test and mid-test data (Research Questions 11-14).

Students from the fourth year of the study (cohort 7) were not included in the analyses for the primary research questions and some of the secondary research questions (i.e., RQ 1-10) because of the forced closure of schools due to COVID-19 in the Spring of 2020 as well as other pandemic-induced changes in students' lives. (For more information, see "Group Changes" in Section III, "Study Design," above.) Therefore, there were only 6 cohorts of 8 schools included in the analytic sample for these research questions, or 48 total clusters (6 cohorts x 8 schools). We counted the last cohort of participants across the 8 schools (1 cohort x 8 schools = 8 clusters) as attrition at the cluster level.

As detailed in Table 8 below, there was a 14% overall attrition rate by cluster and a 9% differential attrition rate by cluster when the students from the fourth year of the study (cohort 7) were excluded for the primary research questions (See Appendix B for CONSORT Diagrams). At the individual level, the maximum attrition rate was 30% (for the pre- to post-test analytic sample) and the maximum differential attrition was 6.6% (for the difference in the percentage of students who took the post-test).

Table 8. Cluster and individual sample sizes by intervention status for primary research questions

	Full Treatment n	Partial Treatment n	Total n	Full Treatment Response Rate	Partial Treatment Response Rate	Total response rate
Clusters						
At beginning of study	20	36	56	n.a.	n.a.	n.a.
Contributed at least one individual at baseline	20	36	56	100%	100%	100%
Contributed at least one individual at post-test	16	32	48	80%	89%	86%
Contributed at least one individual at 3-mo follow-up	16	32	48	80%	89%	86%
Individuals in non-attriting clusters						
At time that clusters were assigned to condition	166	339	505	n.a.	n.a.	n.a.
Who consented	166	338	504	>99%	>99%	100%
Contributed a baseline survey	162	328	490	97%	97%	98%
Contributed to post-test	111	249	360	67%	73%	71%
Contributed to the impact analysis at post-test[a]	109	243	352	65%	72%	69%
Contributed to 3-mo follow-up survey	136	378	414	82%	82%	82%
Contributed to the impact analysis at 3-mo follow up*	134	270	404	81%	80%	80%

Note. n.a. = not applicable.

* accounts for item non-response and any other analysis restrictions. There were 56 clusters in the study, composed of 7 cohorts with 8 schools in each cohort.

[a] accounts for item non-response and any other analysis restrictions. “Intervention” represents students in the full treatment condition. “Control” represents students in the two partial treatment conditions.

B. Outcome measures

Table 9 details the exact measures used to assess program outcomes. To assess coparenting attitudes, we used the Expectations to Coparent scale (Markman, Ganong, & Coleman, 2007). To assess coparenting communication and conflict, we used the Co-parental Communication Scale (Ahrons, 1981). The relationship attitudes scale comes from the healthy relationships (Section A3) scale provided in the HMRF nFORM survey. Next, relationship conflict comes from the Conflict in Adolescent Dating Relationships Inventory (CADRI; Wolfe et al., 2001). To assess depression, we used the Center for Epidemiological Studies Depression Scale (CES-D; Devins & Orme, 1985; Radloff, 1977). To assess worry, we used the Penn State Worry Questionnaire (PSWQ; Beck, Stanley, & Zebb, 1995). All scales have been previously validated and shown to have reliability at or above .70. See Appendix C for the exact survey instrument.

Table 9. Measures related to our primary (RQ 1-4) and secondary analyses (RQ 5-14)

Outcome name	Description of outcome measure (reliability estimates)	Source of the measure	Timing of measure
Coparenting Attitudes (addressing RQ1, RQ5, RQ11)	The outcome measure is a scale (values range from 1 to 4) calculated as the average of 7 survey items measuring general attitudes and beliefs about co-parenting. High scores=more positive attitudes (a ≥ .90).	Local evaluation pre, mid, post, and follow up surveys	At pre-test (week prior to beginning of intervention), at mid-test (week following ending of first semester intervention) post-test (week following ending of second semester intervention), and follow-up (approximately 3 months after the end of intervention)
Coparental Conflict (addressing RQ2, RQ6, RQ12)	The outcome measure is a scale (values range from 1 to 5) calculated as the average of 4 survey items measuring the degree of conflict. High scores=more conflict (a ≥ .85).		
Coparental Interaction (addressing RQ2, RQ6, RQ12)	The outcome measure is a scale (values range from 1 to 5) calculated as the average of 7 survey items measuring how often parents communicate with one another regarding their child. High scores=more communication (a ≥ .90).		
Relationship Attitudes – A (addressing RQ3, RQ7. RQ13)	The outcome measure is a scale (values range from 1 to 4) calculated as the average of 7 survey items measuring general attitudes and beliefs about healthy relationships (e.g., respect, communication, fidelity). High scores=more positive attitudes (a ≥ .76).		
Relationship Attitudes - B (addressing RQ3, RQ7. RQ13)	The outcome measure is a scale (values range from 1 to 4) calculated as the average of 7 survey items measuring general attitudes and beliefs about the use of intimate partner violence in a relationship. High scores=more positive attitudes (a ≥ .78).		
Relationship Conflict - Self (addressing RQ4, RQ8, RQ14)	The outcome measure is a scale (values range from 1 to 4) calculated as the average of 31 survey items measuring respondents' perception of their behaviors during conflict. High scores=more conflict (a ≥ .75).		
Relationship Conflict - Partner (addressing RQ4, RQ8, RQ14)	The outcome measure is a scale (values range from 1 to 4) calculated as the average of 31 survey items measuring respondents' perceptions of their partners' behaviors during conflict. High scores=more conflict (a ≥ .82).		
Depression (addressing RQ9)	The outcome measure is a scale (values range from 1 to 4) calculated as the average of 14 survey items measuring how often participants had experienced symptoms of depression over the last month. High scores = higher depression (a ≥ .88).		
Worry (addressing RQ10)	The outcome measure is a scale (values range from 1 to 5) calculated as the average of 7 survey items measuring participants' feelings of worry. High scores = higher anxiety (a ≥ .95).		

C. Baseline equivalence and sample characteristics

In this section, we assess if there is baseline equivalence across all treatment conditions with regard to demographic characteristics (e.g., age, gender, ethnicity, pregnant vs. parenting status), relationship characteristics (e.g., in a relationship or not, part of a dyad in the study or not), and baseline measures (i.e., pre-test scores) for each analytic sample (Table 10).

Hedges' g (for continuous variables) and Cox (for dichotomous variables) were used to compare treatment groups and assess baseline equivalence, with effect size (ES) <= .|05| indicating equivalence per WWC guidance (What Works Clearinghouse, 2020). If the ES is between |.05| and |.25|, then the corresponding variable(s) are added to the models for each RQ to adjust for the differences statistically. If ES is greater than |.25|, then new groups are constructed utilizing Propensity Score Matching (PSM) techniques.

Table 10. Summary statistics of key baseline measures and baseline equivalence (BE) across study groups, for individuals/couples completing pre- and post-test

Baseline measure		Full Treatment	Partial Treatment	Full Treatment versus Partial Treatment	Effect size
	n	% / Mean (SD)	% / Mean (SD)	Difference (p)	Hedge's g or Cox index
Participant Characteristics					
Female (%)	352	87%	86%	1% (0.77)	<0.05
Pregnant (vs. Parenting) (%)	352	44%	52%	-8% (0.15)	0.19
Hispanic (%)	352	83%	85%	-2% (0.53)	-0.09
Non-Hispanic White (%)	352	8%	6%	2% (0.38)	0.19
Non-Hispanic Black (%)	352	9%	11%	-2% (0.66)	-0.13
In a relationship (%)	352	72%	76%	-4% (0.46)	-0.09
In a participating dyad (%)	352	18%	24%	-6% (0.25)	-0.14
Age	352	16.7 (1.20)	16.7 (1.20)	0% (0.96)	0.01
Birthplace = USA	352	82%	74%	8% (0.16)	0.19
Outcome Measures (Pre-Test)					
Expectations to Coparent (R 1-4)	190	3.36 (0.45)	3.33 (0.58)	0.03 (0.70)	<0.05
Coparenting Conflict (R 1-5)	215	1.98 (0.82)	2.01 (0.81)	-0.03 (0.77)	0.02
Coparenting Interaction (R 1-5)	214	3.98 (1.02)	3.72 (1.10)	0.26 (0.05)	0.24
Healthy Relationship Attitudes A (R 1-4)	327	3.73 (0.34)	3.71 (0.41)	0.02 (0.70)	<0.05
Healthy Relationship Attitudes B (R 1-4)	327	3.41 (0.48)	3.47 (0.48)	-0.06 (0.33)	-0.13
CADRI-Self (R 1-4)	226	1.54 (0.22)	1.54 (0.21)	0.00 (0.89)	0.00
CADRI-Partner (R 1-4)	234	1.59 (0.34)	1.60 (0.27)	-0.01 (0.85)	-0.03

Notes: n.a. = not applicable; p-values are included in parentheses. Effect sizes are calculated using hedge's g (continuous variables) or Cox's index (dichotomous variables). For the participant characteristics section, all participants who provided pre and post-test data were included in the BE analysis noted above. For the outcome measures, individual baseline equivalency tests were run for each outcome variable separately. Each BE analysis included a sample that provided data for the covariates AND the outcome of interest. Therefore, the n next to each outcome represent the analytic sample for the respective BE analysis.

As can be seen in Table 10 above, there were some small differences between groups at baseline. However, there were no variables with ES >.25 at baseline. There were more students who were parenting at pre-test in the full treatment condition than the partial treatment condition. There were slightly fewer Hispanic and Black students and more White students in the full treatment condition. Slightly fewer students in the full treatment condition were in a relationship at pre-test or were part of a participating dyad. Slightly more students in the full treatment condition were

born in the United States than their counterparts. Students in the full treatment condition had a higher existing level of interaction with the other coparent at pre-test and slightly less healthy relationship attitudes. All variables with ES > |0.05| were added to the models examining post-test differences.

We ran similar equivalency analyses for 3-month follow up and mid-test. Tables examining baseline equivalence for these analyses can be found in Appendix B.

V. FINDINGS AND ESTIMATION APPROACH

A. Implementation evaluation

Key Findings:

- Overall, lessons were delivered with fidelity, although adaptations were made to address special requirements due to language abilities, online lesson delivery, or make up lessons.
- When using a 6/10 lesson completion cut off, lesson completion rates were acceptable (> 70%) when assessing completion per curriculum; however, only 44-46% of students reached this threshold across two consecutive semesters.
- Overall, implementation quality and student engagement were high based on program coordinator observation forms and participants' qualitative responses

Table 11 provides a detailed description of each implementation question, the indicator of interest, and final results. These data provided an overview of the fidelity, dosage, quality, engagement, and context of our implementation.

Table 11. Implementation Analysis questions and results

Research question	Results
Were all intended intervention components offered and for the expected duration? (Fidelity; Dosage Provided)	Total number of sessions delivered: • 700 lessons (M = 87.5 lessons/semester) % of observed lessons that met the 75% lesson content cut off: • 56/57 (98%) observed lessons met the 75% content cut off.

Research question	Results
What were the unplanned adaptations to key intervention components, per curriculum? (Fidelity)	**List of unplanned adaptations due to make ups, and online service delivery:** • List of general adaptations: More hands-on activities were created for content; tailoring to developmental stage, specific home experiences, and/or size of group; Spanish translation of all materials; update old/outdated videos (Love Notes); the order of material within a lesson were switched to end on a positive note; and schools that held short lessons (30 minutes) had to have lessons split into two-day lessons. • Adaptations due to online lessons: hands-on activities had to be converted into videos or group discussions (more so in CoP curriculum) • Adaptations due to make-up lessons: lessons were shortened because group discussion could not be completed, and hands-on activities were turned into videos or 1:1 discussions. **% students using online service delivery:** • The % of students attending lessons ranged from 10% (Travis High School) to 85% (Hays High School). On average, 66% of students had attended at least one lesson online before COVID-19. After COVID-19, all students who remained actively engaged, received lessons online. **% of lessons online vs. in-person:** • 672 lessons (96%) were delivered in person and 302 of these in-person lessons (43%) were livestreamed to support students who had to join virtually (43%). 28 lessons (4%) were delivered only in an online format due to COVID-19. % students receiving make-ups: • The % of students who required at least one make-up lesson ranged from 50-60% (Average = 57.6%).
How many students attended the two-intervention model, per intervention assignment? (Dosage Received)	# of students who attended at least one lesson across two consecutive semesters: • 310/610 (51%) students attended at least one lesson across two consecutive semesters. # (%) of students who completed 6/10 and 10/10 lessons per intervention assignment, across two consecutive semesters: • **Full treatment condition (n = 166)**: 77 (46%) students completed at least 6/10 lessons across both semesters, and 25 (15%) completed all 10/10 lessons across both semesters. • **Partial Treatment condition (n = 339):** 149 (44%) students completed at least 6/10 lessons across both semesters, and 51 (15%) completed all 10/10 lessons across both semesters. • Differences in completion rates were 2% when using the 6/10 completion threshold, and 0% when using the 10/10 completion threshold.
How many classes did students attend, on average, per curriculum? (Dosage Received)	% of students who complete 6/10 and 10/10 lessons: • Coparenting curriculum (n = 420): 303 (72%) students completed at least 6/10 lessons and 129 (34%) completed all 10/10 lessons. • Relationships curriculum (n = 414): 301 (73%) students completed at least 6/10 lessons and 143 (35%) completed all 10/10 lessons. • Life After Graduation curriculum (n = 420): 303 (72%) of students completed at least 6/10 lessons and 145 (35%) completed all 10/10 lessons. • Differences in completion rates were 1% across both completion thresholds.

Research question	Results
What was the quality of staff–participant interactions, per curriculum? (Quality)	% of sessions with high quality lessons: • 63/65 (97%) of observed lessons met or exceeded expectations in the classroom layout category (e.g., room is conducive to positive group interaction). • 62/64 (97%) of observed lessons met or exceeded expectations in the quality lesson facilitation category (e.g., loud and clear voice, appropriate speed, lesson comprehension, confidence and control in moderating conversation). Average nFORM responses regarding lesson utility and satisfaction: • Students think SR/SF helped, M = 1.15 (1 = A lot, 2 – Some, 3 = Not at all) • Students learned what makes a healthy relationship (M = 1.46), new skills (M = 1.47), and felt confident in the skills they learned (M = 1.52; 1= Strongly agree to 5 = Strongly disagree) # of positive comments made within participant focus groups: • Positive comments regarding SR/SF was the highest coded theme in our qualitative data. It appeared 3 times more than the second most common theme (151 vs. 49).
How engaged were intervention group members in the intervention, per curriculum? (Engagement)	% of sessions with high student engagement: • 56/58 (97%) of observed lessons met or exceeded expectations in the relationship with participants category. List of qualitative themes that emerge in the nFORM and SR/SF qualitative questions: • Top 5 themes (# of comments per themes): – Positive comments (151), – Enjoyed learning about relationships (49), – Appreciation for the support received from facilitators (26), – SR/SF helped them prepare for the future (18), Building community with other teens (12) • Below are quotes from participants in response to the question "*What was the most valuable thing you learned in SR/SF this year?":* – "The most valuable thing I learned last year was when people shared ways to be a good parent and also practice self-care. For example, it's okay to be stressed but you also need to ask for help if you need it. As a parent you need to be good so that you can be the best parent." – "That a relationship is work and if you want it to work then you have to put a lot of hard effort into it." – "Learning how to listen to others, and using the red, yellow, green light use in a conflict. Not only did they help me stay always from starting a problem, but it helped me in controlling my anger over small situations." – "I learned that there is many ways to communicate with others and that violence is never one of them." – "To stop & think if I'm heated either walk away for a little bit while I cool down then come back to the problem when I'm in a better mood." – "That it's important to be the best parent for your child even if you don't get along with the other parent sometimes. To co-parent and make the best decisions for our child together as much as possible." – "Every single lesson was a piece to the puzzle of learning how to manage and get through conflict, as well as appreciating others and learning how to communicate feelings."

Research question	Results
What external events affected implementation? (Context)	List of external events and # of sites affected by each event: • Hurricane Harvey (Fall 2017) – 8 schools (80 students) were impacted by delayed starts and reduced student attendance in early Fall. *acute impact • Reduction of Austin Independent School District funding/childcare closure (2017-2020) – 4 AISD schools (167 students) showed shifts in student enrollment and student attendance due to an increased need for busing between schools, 2 schools lost childcare centers completely, and one school lost all adolescent parents within a year (5 students). As a result, we left this school and moved to a new school in Fall 2018. *long-term impact • One school within the Comal ISD district changed principals – this led to school policy changes encouraging adolescent parents to transfer to non-traditional schools. We left this school in Fall 2019 and moved to a new school. *long-term impact • COVID-19 (Spring 2020) – 7 schools (76 students) moved to online instruction beginning in March 2020. 5/10 SR/SF lessons moved online. *long-term impact

Summary description of findings. Lesson fidelity was high when assessing the facilitators' administration of lessons and lesson content; however, adaptations had to be made to accommodate participants' needs (e.g., language needs, online lesson attendance, and make-up lessons). Based on PF input, over half of the students attended at least one lesson online or had to complete a make-up lesson. Although dosage provided (# of lessons and lesson content) and dosage received (students attendance and program completion) were acceptable when assessing participant completion using a 6/10 lesson threshold for each curriculum, dosage received was unacceptable when using a 10/10 thresholds (34-35%). Dosage across the two-semester model was also low (44-46%). Quality of implementation appeared to be high across indicators, as was student engagement. Finally, the context of implementation provided insights as to why dosage received, in particular, was a struggle. Several local policy changes, budget cuts, and school leadership changes led to reductions in support services for adolescent parents (e.g., access to childcare), which impacted student attendance and program completion. Further, a global event (COVID-19) impacted the final implementation of our program, which led the removal of a potential study cohort (76 potential study participants) due the severity of the situation.

Limitations. We believe most of the data used were reliable and free of bias; however, we recognize some limitations with the student attendance and adaptation data. Specifically, student attendance had to be calculated by our internal documents instead of nFORM due to the fact that nFORM changed its method of calculation in the middle of the grant cycle, which forced our team to change the manner it tracked attendance. Further, neither nFORM nor our internal tracking sheets had a simple method of noting whether students' attendance or completion of a lesson occurred in person, online, or via a make-up session. For this reason, documentation of variations in lesson delivery received was not available and a retrospective focus group was required. All other data were collected throughout the cycle of the grant.

B. Primary impact evaluation

Key Findings:

- The two recruitment strategies that resulted in the largest share of participants starting When assessing our primary research questions we found there were no statistically ($p < .05$) significant differences between the full and partial treatment groups except on the Coparenting Conflict scale, where students in the full treatment group reported less conflict than their counterparts.
- One other outcome approached significance ($p \leq .10$), such that students in the full, versus partial, treatment group reported slightly healthier relationship attitudes.

We examined program effects in two ways:

T-test Analyses. First, we compared the means of the outcome variables across the full and partial-treatment conditions using t-test analyses. The results from these analyses are considered less reliable because there were some characteristics and pre-test measures for which the groups had an ES > 0.05 and thus did not satisfy the criterion of baseline equivalence. However, they provide an indication of the differences between treatment groups and provide context for the other findings, so they are included in our results tables. They should be considered as preliminary analyses.

Multilevel Models. For each outcome, a multilevel model (MLM) was constructed that clustered students within schools. MLMs take into account that participants within a cluster (e.g., a school) likely share characteristics that make them more similar to each other than to students at other schools. Failing to take these characteristics into account tends to lower the standard errors of regression coefficients, making it more likely to have effects deemed 'significant' when they are not.

Because our models included level-1 (student-level) and covariates, the impact coefficient ($\beta 1j$) measures the net magnitude of the full treatment on student outcomes after controlling for prior score and other school and student factors (See Table 12 below for our modeling approach; See Appendix D for additional data preparation information).

Table 12. Modeling Approach

Two-level Model
Level 1: Yij= β0j + β1 * Treatmentij + B2 * Pretestij + ... + eij Level 2: β0j = γ00 + γ01 * School + μ0j
Where – Y represents the outcome of interest – postscripts i and j index, respectively, student and school – βs and γs are parameters to be estimated; β0j is a random parameter but all other parameters are fixed. – Treatment is a binary indicator (1 if full treatment group, else 0). This variable is treated as a fixed effect. – "…" indicates that the model will include multiple predictors and corresponding parameters – eij represents the individual residual for student i in school j. – us are school-specific residuals (estimated as random effects) and they are independently and identically distributed with a mean of 0.

Models were estimated using the R package 'lme4'. We used the standard alpha level of $p < .05$ as an indicator of statistical significance.

Covariates. Table 13, below, lists the covariates used in the impact analyses. These covariates had a baseline effect size greater than 0.05. An additional item, the frequency of interaction with the other biological parent, was added to the coparenting analyses. We reasoned that adolescent parents who interacted more with their child's other biological parent might have different expectations about coparenting, and different communication styles than adolescent parents who interacted less often with their child's other parent.

Table 13. Covariates included in impact analyses

Covariate	Description of the covariate
Ethnicity: Hispanic	Self-identifies as Hispanic (1 = Hispanic; 0 = not Hispanic)
Race: White	Self-identifies as White (1= White; 0 = not White)
Race: Black	Self-identifies as Black (1= Black; 0 = not Black)
Pregnant	Pregnant at pre-test (1=pregnant, 0 = currently parenting)
Romantic Partner	Has romantic partner at pre-test (1=has partner, 0=no partner)
Dyad	Part of a participating dyad in SR/SF (1=dyad, 0 = participated alone)
Birthplace	Place of birth (1 = USA; 0 = Outside USA)
CP Interaction Pre-Test	Coparenting Interaction – Other Biological Parent – pre-test scale mean
Healthy Relationship Pre-Test	Healthy Relationship Attitudes – B – pre-test scale mean
Frequency (RQ 1-2)	Frequency of interaction with other parent pre-test (0=none to 6=daily) - added to analyses for RQ 1 and 2 only

Notes. See Table 9 for a more detailed description of each measure. Additionally, the student's pre-test mean was added to the analysis for that outcome variable (e.g., Expectations to Coparent pre-test score was added to the analysis of Expectations to Coparent at post-test)

Intent to treat. We used an intent-to-treat approach. Therefore, we estimated the effect of being assigned to the treatment condition, regardless of whether they completed the treatment condition or were transferred into another school (and received another treatment condition). This was a rare occurrence – only two students in the analytic sample changed schools and conditions.

Results for the Primary Research Questions

Key results are presented in Table 14 and summarized below. For detailed information on analyses, with values for all models, see Appendix E.

RQ1. There were no differences between the full and partial treatment groups, either in the t-test analyses or in the multilevel models in terms of coparenting attitudes. The mean for the scale was 3.4 (of 4), indicating that students in both groups had fairly positive attitudes.

RQ2. In the multilevel model, results indicated that students in the full treatment condition had lower conflict with their child's other parent than students in the partial treatment condition (B = -0.29, p <.05; See Appendix E for complete MLM results).

There were no differences between conditions in terms of coparenting interaction. Both groups reported interacting with their child's other parent fairly frequently (M = 3.98 on a scale of 5).

RQ3. There was a marginally significant effect of condition on students' relationship attitudes on the first scale administered (HRA). Students in the full treatment condition had slightly healthier relationship attitudes than students in the partial treatment condition (M = 3.79 vs. 3.71 on a scale of 4, p = .09). However, there was not an effect in the MLM once pre-test score and other participant characteristics were controlled.

There was not a significant effect of condition on students' relationship attitudes on the second scale (HRB). On average, students demonstrated very healthy relationship attitudes on both relationship attitude scales (Overall *M* = 3.74 of 4 and *M*=3.41 of 4).

RQ4. There was not a significant effect of condition on students' reported relationship conflict. Both reported low scores for negative behaviors exhibited either for themselves (*M* = 1.54 of 4) or their partners (*M* = 1.60 of 4).

Table 14. Summary of Findings for Primary Research Questions (Post Test Differences Between Groups)

Outcome measure	Full Trt	Partial Trt	Full Trt Post-Test	Partial Trt Post-Test	Mean difference	Effect size	MLM Coefficient
	n	*n*	*M(SD)*	*M(SD)*	*Δ (p)*	*g*	*γ1 (p)*
Expectations to Coparent (RQ1)	62	171	3.41 (0.64)	3.40 (0.66)	0.01 (0.92)	0.02	-0.14 (0.27)
Coparenting Conflict (RQ2)	87	149	1.89 (0.82)	2.09 (0.85)	-0.20 (0.07†)	-0.24	-0.28 (0.03*)
Coparenting Interaction (RQ2)	88	148	4.00 (1.11)	3.96 (1.04)	0.04 (0.80)	0.04	0.08 (0.59)
Healthy Relationship Attitudes A (RQ3)	99	227	3.79 (0.33)	3.71 (0.41)	0.08 (0.09†)	0.21	0.59 (0.26)
Healthy Relationship Attitudes B (RQ3)	99	227	3.45 (0.52)	3.39 (0.59)	0.06 (0.30)	0.11	0.06 (0.40)
CADRI – Self (RQ4)	84	177	1.57 (0.34)	1.53 (0.27)	0.04 (0.34)	0.14	0.05 (0.22)
CADRI – Partner (RQ4)	83	174	1.62 (0.52)	1.59 (0.31)	0.03 (0.55)	0.09	0.07 (0.15)

Source: SR/SF Post-test survey and multilevel statistical model

* significant at p < .05;

† marginally significant at p < .10.

Trt = Treatment. Model coefficients presented are from multilevel models controlling for students' pre-test scores as well as all of the items in Table 13.

C. Sensitivity analyses

Key Findings:

Students from the fourth year of the program were removed from analyses of the primary research questions because of the changes to the program (and students' lives) caused by the COVID-19 pandemic. We conducted sensitivity analyses by including these students in the MLMs, plus a new variable indicating if the student had been in the group affected by COVID-19 (0=no, 1=yes).

Generally, we found the same pattern of results as in the previous section; that is, there were no significant differences between groups except for on the Coparenting Conflict scale, where students in the full treatment group reported less conflict than their counterparts. There was an additional improvement for the full treatment group in terms of their healthy relationship attitudes, which was significant for the t-test analysis and marginally significant in the MLM

Sample. Students who participated in the final year of the SR/SF program were heavily impacted by the COVID-19 pandemic in the spring of 2020. All schools closed, and the program was forced to be conducted virtually. Because of the changes to the program and to students' lives, these students were excluded from the sample examined in the primary research questions. The sensitivity analyses explored what differences emerged when adding them back into the sample.

Covariates. Table 15, below, contains the covariates used in the sensitivity analyses (See Appendix F for more details). These covariates had a baseline effect size greater than 0.05. An additional item, the frequency of interaction with the other biological parent, was added to the coparenting analyses, and the presence of a romantic partner was added to the healthy relationship analyses.

Table 15. Covariates included in impact analyses

Covariate	Description of the covariate
Race: Black	Self-identifies as Black (1= Black; 0 = not Black)
Pregnant	Pregnant at pre-test? (1=pregnant, 0 = currently parenting)
Romantic Partner (RQ 3-4)	Has romantic partner at pre-test (1=has partner, 0=no partner)
Birthplace	Place of birth (1 = USA; 0 = Outside USA)
CP Conflict Pre-Test	Coparenting Conflict – Other Biological Parent – pre-test scale mean
CP Interaction Pre-Test	Coparenting Interaction – Other Biological Parent – pre-test scale mean
Healthy Relationship B Pre-Test	Healthy Relationship Attitudes – B – pre-test scale mean
Frequency (RQ 1-2)	Frequency of interaction with other parent pre-test (0=none to 6=daily) - added to analyses for RQ 1 and 2 only

Notes. See Table 9 for a more detailed description of each measure. Additionally, the student's pre-test mean was added to the analysis for that outcome variable (e.g., Expectations to Coparent pre-test score was added to the analysis of Expectations to Coparent at post-test)

Findings. Generally, we saw the same pattern of results as reported in the "Primary Impact Evaluation" section above (Table 16). Students in the full treatment condition were less likely to report conflict with their child's other parent, t (276) = 1.95, p < .05. This finding held in the covariate MLM model. Additionally, students in the full treatment condition had healthier relationship attitudes that their counterparts (t (376) = 2.63, p < .01). In the covariate MLM, this finding was only marginally significant.

Table 16. Summary of Findings for Primary Research Questions (Post Test Differences Between Groups) – Sensitivity Analyses using students from all cohorts

Outcome measure	Full Trt n	Partial Trt n	Full Trt Post-Test M(SD)	Partial Trt Post-Test M(SD)	Mean difference Δ (p)	Effect size g	MLM Coefficient γ1 (p)
Expectations to Coparent (RQ1)	89	193	3.44 (0.60)	3.38 (0.65)	0.06 (0.45)	0.09	-0.09 (0.3)
Coparenting Conflict (RQ2)	109	168	1.86 (0.79)	2.06 (0.83)	-0.20 (0.05*)	-0.25	-0.24 (0.02*)
Coparenting Interaction (RQ2)	109	167	4.06 (1.10)	3.96 (1.01)	0.10 (0.40)	0.10	0.08 (0.55)
Healthy Relationship Attitudes A (RQ3)	127	250	3.81 (0.30)	3.72 (0.41)	0.09 (<0.01*)	0.24	0.08 (0.08) †
Healthy Relationship Attitudes B (RQ3)	127	250	3.50 (0.51)	3.41 (0.58)	0.09 (0.11)	0.16	0.01 (0.85)
CADRI – Self (RQ4)	107	192	1.54 (0.32)	1.53 (0.27)	0.01 (0.62)	0.03	0.04 (0.29)
CADRI – Partner (RQ4)	106	189	1.60 (0.36)	1.59 (0.30)	0.01 (0.81)	0.03	0.05 (0.28)

Source: SR/SF Post-test survey and multilevel statistical model.

* significant at p < .05;

† marginally significant at p < .10.

Trt = Treatment. Model coefficients presented are from multilevel models controlling for students' pre-test scores and COVID status as well as all of the items in Table 18.

D. Additional analyses

1. Secondary Research Questions

Key Findings:

- When assessing our secondary research questions regarding group differences at three-month follow up, no differences emerged across treatment groups.
- When comparing participants who completed the coparenting vs. control curriculum, no statistical differences emerged. However, one exploratory analyses approached significance ($p \leq .10$); specifically, students who attended the coparenting curriculum reported slightly more communication with the other biological parent of their child at mid-test. No differences emerged when assessing outcome differences between students attending the romantic relationship curriculum vs. control.

Follow Up Research Questions (RQ 5-10). The next set of research questions examined the same outcomes at 3-month follow-up, and also added two new measures: depression and anxiety. There were no differences between conditions at 3-month follow up for any of the research questions, either in the t-tests or the MLMs. Results are summarized in Table 17. For detailed information on analyses, with values for all models, see Appendix E.

Table 17. Summary of Findings for Secondary Research Questions 5-10 – Assessing 3-month Follow up Differences

Outcome measure	Full Trt n	Partial Trt n	Full Trt Post-Test M(SD)	Partial Trt Post-Test M(SD)	Mean difference Δ (p)	Effect size g	MLM Coefficient β (p)
Expectations to Coparent (RQ5)	98	211	3.36 (0.60)	3.38 (0.62)	-0.02 (0.81)	-0.03	0.02 (0.87)
Coparenting Conflict (RQ6)	80	168	1.87 (0.85)	1.97 (0.86)	-0.10 (0.39)	-0.12	-0.20 (0.12)
Coparenting Interaction (RQ6)	80	168	4.20 (1.01)	4.16 (1.06)	0.04 (0.78)	0.04	0.11 (0.42)
Healthy Relationship Attitudes A (RQ7)	99	217	3.63 (0.67)	3.73 (0.53)	-0.10 (0.19)	-0.17	-0.12 (0.16)
Healthy Relationship Attitudes B (RQ7)	99	216	3.51 (0.55)	3.42 (0.63)	0.09 (0.20)	0.15	0.04 (0.62)
CADRI – Self (RQ8)	73	163	1.50 (0.29)	1.50 (0.28)	0.00 (0.99)	0.00	-0.02 (0.67)
CADRI – Partner (RQ8)	73	163	1.56 (0.28)	1.58 (0.27)	-0.02 (0.50)	-0.07	-0.02 (0.62)
Depression (RQ9)	100	214	1.64 (0.63)	1.63 (0.55)	0.01 (0.83)	0.02	0.05 (0.53)
Anxiety (RQ10)	100	215	2.27 (1.24)	2.18 (1.10)	0.09 (0.57)	0.08	0.04 (0.81)

Source: SR/SF Post-test survey and multilevel statistical model

* significant at $p < .05$;

† marginally significant at $p < .10$.

Trt = Treatment. Model coefficients presented are from multilevel models controlling for students' pre-test scores as well as all of the items in Table 13.

Mid-Test Research Questions (RQ 11-14). The final set of research questions used slightly different comparison groups. Research Questions 11 and 12 compared the coparenting attitudes, conflict, and interaction outcomes for students who completed the coparenting curriculum and those who completed the control curriculum. Research Questions 13 and 14 compared the romantic relationship attitudes, conflict, and interaction outcomes for students who completed the healthy relationship curriculum and those who completed the control curriculum. Results are summarized in Tables 18 and 19. For detailed information on analyses, including values for all models, see Appendix E.

Coparenting vs. Control. There were no significant differences between students in the coparenting condition and those in the control condition on coparenting attitudes or communication at mid-test (Table 18). There was one marginal finding for the t-test for coparenting interaction (t (207) = 1.63, p = .10.) Students receiving the coparenting curriculum were marginally more likely to report higher interaction levels with their child's other parent.

Table 18. Summary of Findings for Secondary Research Questions 11-12 - Comparing the Coparenting (COP) Group to the Control (CONT) Group at Mid-Test

Outcome measure	CoP Group n	CONT Group n	CoP Post-Test M(SD)	CONT Post-Test M(SD)	Mean difference Δ (p)	Effect size g	MLM Coefficient B1 (p)
Expectations to Coparent (RQ11)	76	70	3.49 (0.47)	3.46 (0.50)	0.03 (0.70)	0.06	-0.03 (0.74)
Coparenting Conflict (RQ12)	98	112	1.99 (0.75)	1.93 (0.84)	0.06 (0.59)	0.08	0.21 (0.07†)
Coparenting Interaction (RQ12)	111	97	4.10 (1.05)	3.86 (1.05)	0.24 (0.10†)	0.23	0.02 (0.89)

Source: SR/SF Post-test survey and multilevel statistical model

* significant at p < .05;

† marginally significant at p < .10. Model coefficients presented are from multilevel models controlling for students' pre-test scores as well as all of the items in Table 13.

Love Notes (Healthy Relationships) vs. Control. There were no significant differences between students receiving the Love Notes curriculum and those receiving the control curriculum in terms of relationship attitudes or relationship conflict at mid-test (See Table 19).

Table 19. Summary of Findings for Secondary Research Questions 13-14 - Comparing the Healthy Relationships (HRR) Group to the Control (CONT) Group at Mid-Test

Outcome measure	HRR n	CONT n	HRR Post-Test M(SD)	CONT Post-Test M(SD)	Mean difference Δ (p)	Effect size g	MLM Coefficient B1 (p)
Healthy Relationship Attitudes A (RQ13)	87	70	3.77 (0.35)	3.64 (0.33)	0.13 (0.14)	0.21	0.13 (0.22)
Healthy Relationship Attitudes B (RQ13)	87	70	3.50 (0.61)	3.63 (0.42)	-0.13 (0.14)	0.24	-0.08 (0.44)
CADRI – Self (RQ14)	87	101	1.50 (0.29)	1.53 (0.28)	-0.03 (0.51)	0.11	-0.01 (0.79)
CADRI – Partner (RQ14)	86	99	1.59 (0.38)	1.57 (0.30)	0.02 (0.66)	0.06	0.01 (0.77)

Source: SR/SF Post-test survey and multilevel statistical model

* significant at p < .05;

† marginally significant at p < .10. Model coefficients presented are from multilevel models controlling for students' pre-test scores as well as all of the items in Table 13.

VI. DISCUSSION

The purpose of this study was to assess the impact of a relationship-based intervention program serving pregnant and parenting adolescents in Central Texas. Although our study first began as a proposed RCT, low enrollment numbers at certain schools required us to change the study design to a QED. Nonetheless, this study benefitted from a large sample of pregnant and parenting adolescents, a hard to reach population, and data across four time-points.

Results from our implementation analyses showed that, overall, the implementation of the lessons was delivered with high fidelity and students were engaged with the lesson content. However, adaptations were required to respond to adolescents' English fluency, barriers to school attendance, and school retention. Many students did not comp phy relationships curricula shared key elements that led to overlap. When students from each condition participated in focus groups, they tended to emphasize the improvements they had made in both their romantic and coparenting relationships, regardless of condition. Additionally, many students were in a romantic relationship with their child's other parent. This overlap likely also diminished our ability to see an impact of the full treatment condition.

Sample size and clustering effects at the school level may have impacted our power to detect effects. As can be seen in tables 14-17, several outcomes showed small effect sizes in the expected directions that may have been significant if sufficiently powered. Because a cohort of students was not included in our analyses due to the impacts of COVID-19, our sample size was reduced; thus, lowering our power.

Combining both the coparenting and control and healthy relationships and control into one condition (the partial treatment condition) may have also damped our ability to see differences between groups. That is, had we analyzed data for three conditions (full treatment, partial treatment: healthy relationships only, partial treatment: coparenting only), we may have found some significant effects.

Many students reported high scores on the healthy attitudes and behaviors scales at pre-test which may have led to floor effects for negative outcomes (e.g., intimate partner violence, conflict) and ceiling effects for positive outcomes (e.g., communication, healthy relationship attitudes). These pre-test results indicate that the chosen measures may not have been sensitive enough to assess differences across participants.

Additionally, because our study relied on self-reported data where participants reported on their own behaviors using survey data, as opposed to reporting on their coparents/romantic partners' behaviors, we expect social desirability effects emerged (i.e., participants reported themselves in a more positive light). Past impact analyses of similar programs that used observational assessments (Feinberg et al., 2010), assessments of partners' behaviors (when assessing dyad outcomes), and/or retrospective survey measures have shown strong program effects. These alternate designs should be considered in future studies.

Past research has suggested that impacts for healthy relationship interventions (HRI) are often small for RCTs (range .30-.40; Hawkins et al., 2008) and even smaller for QEDs; therefore, our results seem to align with past research. Researchers suggest that these small effects may be, for example, because these types of curricula often increase participants' awareness of their own – and their partners' – negative behavior. Such participants may have rated themselves positively prior to starting the HRI using less stringent criteria, and more negatively as they learn about the nuances of such behaviors after completing the HRIs. It is also quite possible that the discussions elicited by HRIs may cause initial discomfort as individuals and dyads may be discussing difficult topics for the first time, leading to conflict that may mask potential program effects. Lastly, it is important to note that few evaluations of HRI programs have been conducted with historically underrepresented populations (Hawkins et al., 2008); in fact, much of the relationship research that informs the development of these interventions are among white, middle class, and intact family constellations (Letiecq, 2019). Therefore, there is a need to assess the cultural appropriateness and responsiveness of the content and evaluation methods used to assess HRIs.

Lessons Learned

There were many valuable lessons learned through this five-year project that are applicable to future work.

One semester each of coparenting and healthy relationship training significantly reduced reported coparenting conflict. Students in the full treatment condition were less likely, for example, to say that conversations about their child with their child's other biological parent were stressful and tense. However, having an additional semester of healthy relationship training appears to have improved the skills even more. That is, having two semesters to practice conflict resolution and communication skills in two different contexts (i.e., parenting and romantic relationships) may provide additional benefits to students than one semester and one context.

Implementing relationship and parenting curricula in traditional schools comes with a mix of benefits and challenges. Serving adolescents at the schools they attend has many positive benefits. Students were in a familiar location and were able to form a stable peer group of other supportive adolescents who were also parenting. By providing groups within the schools and during the school day, we also remove the burden of having to attend lessons after school and prevented the need to add one more task to adolescents' already busy schedules.

On the other hand, there are unique challenges to serving students in schools. Programs must be flexible enough to address diverse locations, group times, and the availability of resources to reduce barriers to participation. Most of the groups were scheduled during lunch, which presented a challenge in that many students had an attractive alternative to group participation (i.e., eating lunch with their friends). Additionally, programs serving adolescent parents in schools are impacted by the high level of mobility of these students, who often drop out or leave traditional schools for more flexible, non-traditional alternatives. Pursuing partnerships with

these schools is an important step, as is balancing the need for fidelity with the need for flexibility to serve all students.

Flexibility is key in getting full participation from adolescent parents. Adolescent parents face a wide array of challenges that impact their ability to participate in a group consistently. They are more likely to miss school due to their own or their children's illnesses. Mothers who give birth while participating in a program will need to discontinue in-person participation for a time, and those with young babies may need to bring their children to group.

The SR/SF team noticed that a large number of adolescents were unable to complete 6/10 lessons in any given semester. The team worked to create a synchronous delivery model, either by phone or online, so that students could participate at home. They also provided one-on-one make-up lessons to students who needed them. These changes required programmatic adaptations to fit the new delivery models and highlight, once again, the need to find a balance between cultural and contextual responsiveness and fidelity.

Providing incentives helps increase lesson completion. SR/SF made an additional innovation to its delivery model after noticing that satisfactory completion rates were low. They piloted a model that provided an attractive monetary incentive to any student who completed 10/10 lessons with no make-ups and a small incentive for completing the majority of lessons with a few make-ups allowed. This model was so successful that it was implemented in all of the schools after one semester. The incentive, and the accountability that it helped to create, changed student behavior. They began to prioritize the program and attending group, or at least complete make-ups. The average number of lessons that students completed increased, for example, from 5.6 in the Fall of 2018 to 8.0 in the Fall of 2019 (Hutson et. al., 2019).

Shorter participation requirements are better for adolescent parents. As mentioned above, students served by SR/SF were highly mobile. Many did not attend their schools for two consecutive semesters, either because they gave birth and left campus for a semester, or because they decided to leave their schools to enroll in alternative, more flexible schools, or to drop out to work or focus on parenting. Thus, even though the two-semester model may have provided some benefits to students, those benefits are likely outweighed by the difficulty in participating for a long period of time. Models that are one semester – or shorter – may ultimately be more beneficial for students. Additionally, virtual participation models may provide opportunities for students who leave school to continue participating with their peers.

Specialized populations present unique challenges. Fathers were one group that was particularly hard to serve. First, it was difficult to recruit fathers, partially because they were not identified in school records. They also sometimes attended different schools than the mothers, which caused issues if the schools were implementing different curricula. Because there were small numbers of fathers who were enrolled in the program, they were likely disincentivized from participating fully.

Spanish-speaking students were also a minority in SR/SF groups, which led to delivery of English classes with live translation and bilingual supplements. Some students reported

frustration with the translation (as it slowed down the pace of the lessons considerably). The gap also made it hard for groups with Spanish speaking students to feel fully cohesive and may have disincentivized students who primarily spoke Spanish from participating.

It is important to find ways to link students from these 'specialized' groups together to form a support network. The virtual model is one promising method. Further, because these groups require unique recruitment and service delivery methods, it may be best to offer specialized services to these groups.

What lessons did students learn? As suggested by our implementation analyses, our participants valued the program, its service delivery model, and the content of the lessons. Positive comments were offered in focus groups and post-test and follow up surveys. For many participants, each curriculum offered something helpful to strengthen their family. Some participants learned about communication and conflict management ("there [are] many ways to communicate with others and that violence is never one of them"), self-care ("it's okay to be stressed but you also need to ask for help if you need it"), or collaboration ("That it's important to be the best parent for your child even if you don't get along with the other parent sometimes. To co-parent and make the best decisions for our child together as much as possible"). As one participant said "Every single lesson was a piece to the puzzle."

VII. REFERENCES

Ahrons, C. R. (1981). The continuing coparental relationship between divorced spouses. American *Journal of Orthopsychiatry*, 51(3), 415-428.

Barbee, A. P., Cunningham, M. R., van Zyl, M. A., Antle, B. F., & Langley, C. N. (2016). Impact of two adolescent pregnancy prevention interventions on risky sexual behavior: A three-arm cluster randomized control trial. *American Journal of Public Health*, 106(S1), S85-S90.

Beck, J. G., Stanley, M. A., & Zebb, B. J. (1995). Psychometric properties of the Penn State Worry Questionnaire. *Journal of Clinical Geropsychology*, 1, 33–42.

Carlson, M. J., McLanahan, S. S., & Brooks-Gunn, J. (2008). Coparenting and nonresident fathers' involvement with young children after a nonmarital birth. *Demography*, 45(2), 461-488.

Centers for Disease Control and Prevention (2018). About teen pregnancy. Available at: https://www.cdc.gov/teenpregnancy/about/index.htm.

Devins, G.M., Orme, C.M. (1985). Center for epidemiologic studies depression scale. In D.J. Keyser and R.C. Sweetland (Eds.), Test critiques (pp. 144-160). Kansas City, MO: Test Corporation of America.

Feinberg, M. E., & Kan, M. L. (2008). Establishing family foundations: Intervention effects on coparenting, parent/infant well-being, and parent-child relations. *Journal of Family Psychology*, 22(2), 253-263.

Feinberg, M. E., Jones, D. E., Kan, M. L., & Goslin, M. C. (2010). Effects of family foundations on parents and children: 3.5 years after baseline. *Journal of Family Psychology*, 24(5), 532-542.

Feinberg, M. E., Jones, D. E., Hostetler, M. L., Roettger, M. E., Paul, I. M., & Ehrenthal, D. B. (2016). Couple-focused prevention at the transition to parenthood, a randomized trial: Effects on coparenting, parenting, family violence, and parent and child adjustment. *Prevention Science,* 17(6), 751-764.

Hawkins, A. J., Blanchard, V. L., Baldwin, S. A., & Fawcett, E. B. (2008). Does marriage and relationship education work? A meta-analytic study. *Journal of Consulting and Clinical Psychology*, 76(5), 723-734.

Hoffman, S. D., & Maynard, R. A. (2008). Kids having kids: Economic costs and social consequences of teen pregnancy (2nd ed). Washington, DC: Urban Institute Press.

Hohmann-Marriott, B. (2011). Coparenting and father involvement in married and unmarried coresident couples. *Journal of Marriage and Family*, 73(1), 296-309.

Hutson, A., Perez-Brena, N. J., Perez, R., Reyes, E., Toews, M., & Zambrano, P (August, 2019). An incentive-based approach to improving program attendance. Paper presented at the Healthy Marriage and Responsible Fatherhood Biennial Conference at National Harbor, MD.

Kerpelman, J. (2010). The YouthBuild USA evaluation study of Love Notes. Retrieved from: http://www.dibbleinstitute.org/Documents/YBUSA-Love-Notes-Evaluation-report-2010.pdf.

Lai, M. H., & Kwok, O. M. (2015). Examining the rule of thumb of not using multilevel modeling: The "design effect smaller than two" rule. *The Journal of Experimental Education*, 83(3), 423-438.

Letiecq, B. L. (2019). Surfacing family privilege and supremacy in family science: Toward justice for all. *Journal of Family Theory & Review*, 11(3), 398-411.

Lewin, A., Mitchell, S. J., Beers, L. S., Feinberg, M. E., & Minkovitz, C. S. (2012). Coparenting in teen mothers and their children's fathers: Evidence from the early childhood longitudinal study–birth cohort. *Academic Pediatrics*, 12(6), 539-545.

Markman, M. S., Ganong, L. H., & Coleman, M. (2007). Coparental identity and mothers' cooperation in coparental relationships. *Family Relations*, 56(4), 369–377.

Martin, J. A., Hamilton, B. E., Osterman, M. J. K., Driscoll, A. K., & Drake, P. (2018). Births: Final data for 2017. National Vital Statistics Report, 67(8). Hyattsville, MD: National Center for Health Statistics. Retrieved from: https://www.cdc.gov/nchs/data/nvsr/nvsr67/nvsr67_08-508.pdf.

Radloff, L. (1977). The CES-D Scale A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 7, 385-401.

Sedgh, G., Finer, L. B., Bankole, A., Eilers, M. A., & Singh, S. (2015). Adolescent pregnancy, birth, and abortion rates across countries: Levels and recent trends. *Journal of Adolescent Health*, 56(2), 223-230.

Texas Department of State Health Statistics (2015). Table 11 Births to Mothers 17 Years of Age and Younger by County of Residence. Retrieved from: https://www.dshs.texas.gov/chs/vstat/vs15/t11.aspx.

Toews, M. L., & Yazedjian, A. (2010). "I learned the bad things I'm doing": Adolescent mothers' perceptions of a relationship education program. Marriage & Family Review, 46(3), 207-223.

Toews, M. L., Yazedjian, A., & Jorgensen, D. (2011). "I haven't done nothin' crazy lately:" Conflict resolution strategies in adolescent mothers' dating relationships. Children and Youth Services Review, 33(1), 180-186.

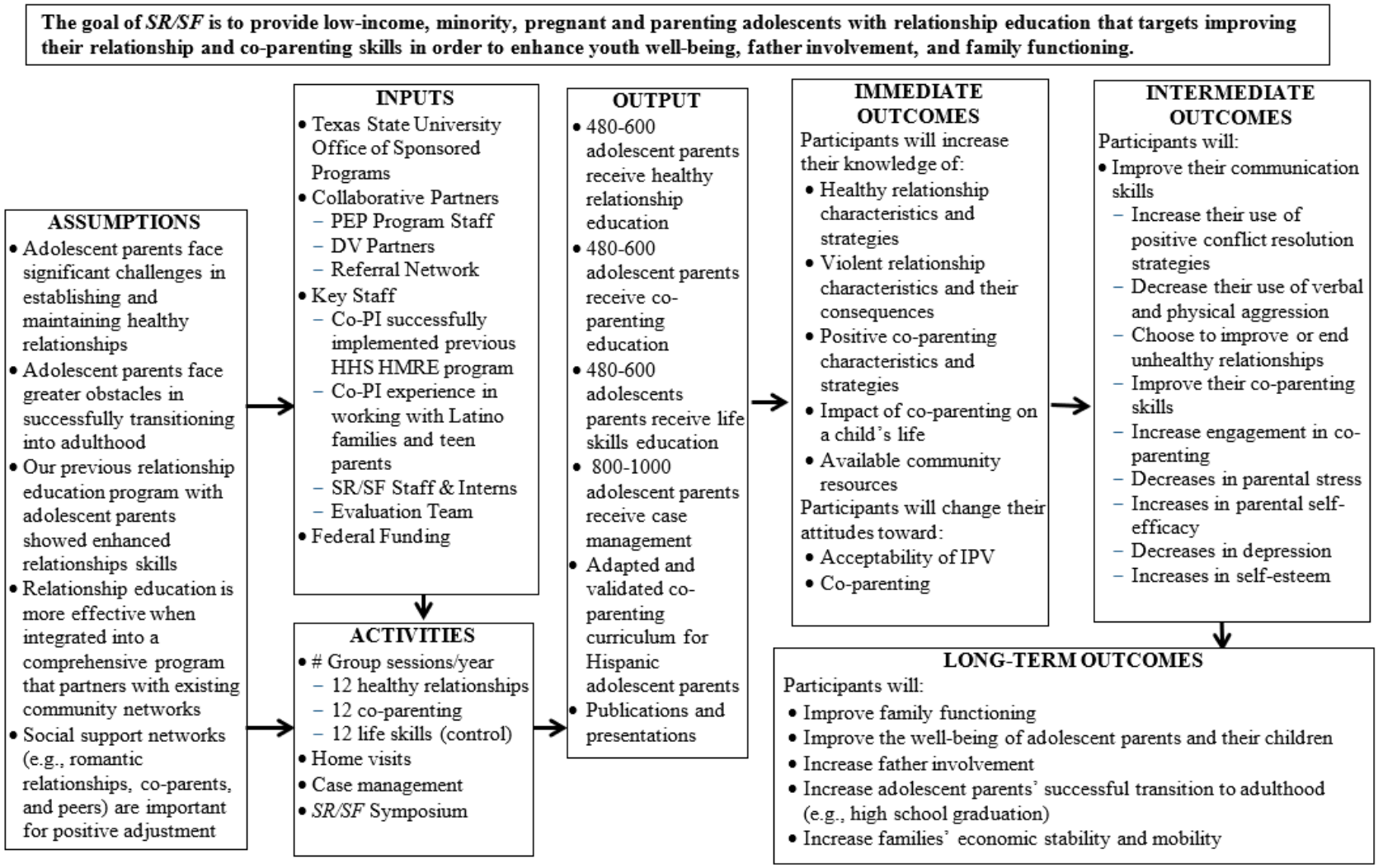
Waite, L. J., & Gallagher, M. (2000). The case for marriage: Why married people are happier, healthier, and better off financially. New York, NY: Doubleday.

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

Wolfe, D. A., Scott, K., Reitzel-Jaffe, D., Wekerle, C., Grasley, C., & Straatman, A. L. (2001). Development and validation of the conflict in adolescent dating relationships inventory. Psychological Assessment, 13(2), 277.

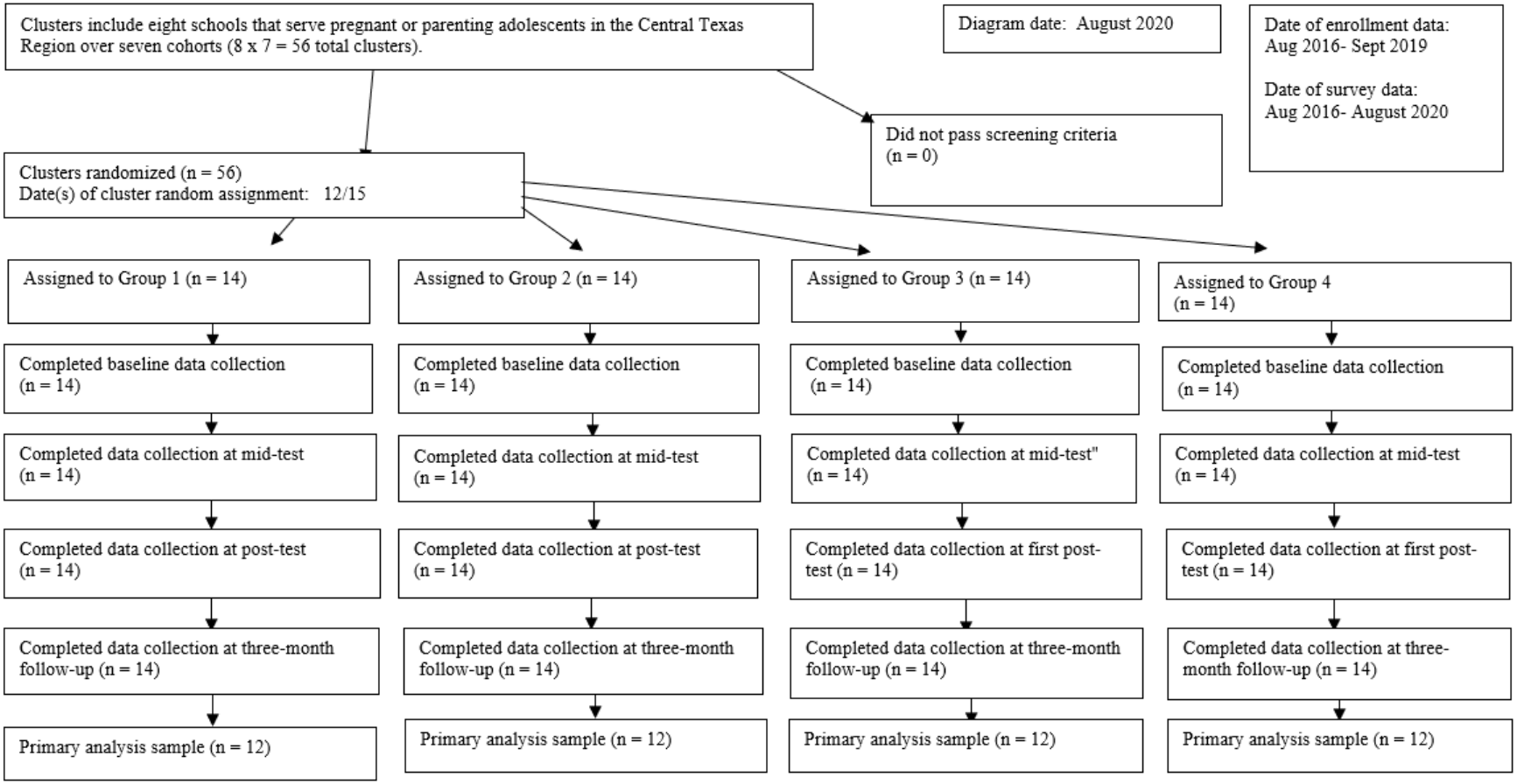
VIII. APPENDICES

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



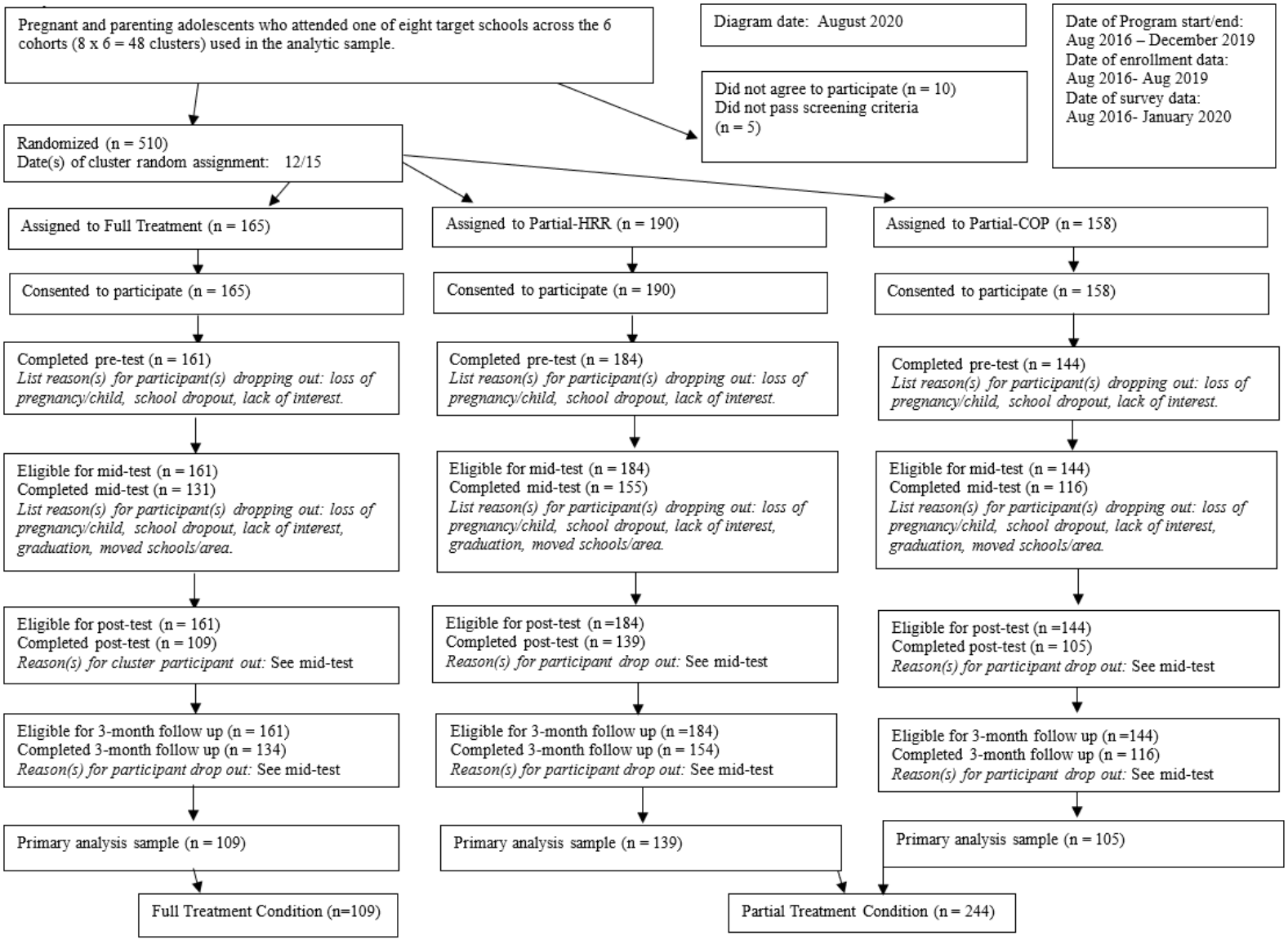
Appendix B. Data and Study Sample.

Figure B.1. CONSORT diagram for Clusters*



*Note. In 2018 and again in 2019, a school in our original clusters was replaced. The n for the primary analysis sample is lower than the total n for each cluster because data for students from the last cohort was removed from the study due to the substantial impact of COVID-19 at the end of the 2019-2020 school year.

Figure B.2. CONSORT diagram for Individual Participants where random assignment occurred before consent



Attrition Rates and Baseline Equivalence of the RCT Design.

For the primary research questions, because our sample showed baseline equivalence across all study variables (See Table 10), we did not need to construct equivalent groups using propensity score matching. Therefore, our final groups reflect what would have been estimated if our study had been an RCT.

In this section, we present the baseline equivalence rates for the secondary research questions which examine differences between students at 3-month follow up and mid-test. Table B.1, below, shows the effect size of the baseline differences for the pre- and 3-month follow up samples. Importantly, there was one difference that crossed the ES >0.25 threshold. Students in the full treatment condition were less likely to be in a participating dyad than students in the partial treatment condition (16% vs 23%, ES = 0.27).

Table B.1. Summary statistics of key baseline measures and baseline equivalence across study groups, for individuals/couples completing pre-test and 3-month follow up

Baseline measure	n	Full Treatment % / Mean (SD)	Partial Treatment % / Mean (SD)	Full Treatment versus Partial Treatment Difference (p)	Effect size Hedge's g or Cox index
Participant Characteristics					
Female (%)	321	88%	86%	2% (0.65)	0.11
Pregnant (vs. Parenting) (%)	321	43%	52%	-9% (0.14)	-0.22
Hispanic (%)	321	82%	86%	-4% (0.33)	-0.18
Non-Hispanic White (%)	321	8%	6%	2% (0.56)	0.19
Non-Hispanic Black (%)	321	10%	10%	0% (0.94)	0.00
In a relationship (%)	321	72%	77%	5% (0.35)	-0.16
In a participating dyad (%)	321	16%	23%	7% (0.14)	-0.27
Age	321	16.7 (1.20)	16.7 (1.20)	0% (0.99)	0.03
Birthplace = USA	321	81%	74%	7% (0.23)	0.16
Outcome Measures (Pre-Test)					
Expectations to Coparent (R 1-4)	169	3.33 (0.44)	3.36 (0.56)	-0.03 (0.72)	-0.06
Coparenting Conflict (R 1-5)	221	3.98 (0.77)	3.83 (0.76)	0.15 (0.15)	0.20
Coparenting Interaction (R 1-5)	221	3.72 (0.34)	3.73 (0.37)	-0.01 (0.87)	-0.03
Healthy Relationship Attitudes A (R 1-4)	315	3.41 (0.48)	3.47 (0.49)	-0.06 (0.28)	-0.22
Healthy Relationship Attitudes B (R 1-4)	314	0.43 (0.50)	0.52 (0.50)	-0.09 (0.14)	0.00
CADRI-Self (R 1-4)	205	1.54 (0.23)	1.54 (0.21)	0.00 (0.99)	0.00
CADRI-Partner (R 1-4)	201	3.33 (0.44)	3.36 (0.56)	-0.03 (0.72)	-0.06

Notes: n.a. = not applicable; p-values are included in parentheses. Effect sizes are calculated using hedge's g (continuous variables) or Cox's index (dichotomous variables). Individual baseline equivalency tests were run for each outcome variable which included all covariates AND the outcome of interest. The n next to each outcome indicates the sample size used for each individual baseline equivalence analyses. Note that "In a participating dyad" is over the baseline equivalence threshold of 0.25. Therefore, these samples cannot be considered equivalent and all results must be interpreted as exploratory.

Table B.2, below, shows the effect size of the baseline differences for the pre- and mid-test sample for the coparenting vs. control samples. There was one difference that crossed the ES >0.25 threshold. Students in the coparenting condition were more likely to be in pregnant at pre-test than students in the control condition (57% vs 43%, ES = 0.34).

Table B.2. Summary statistics of key baseline measures and baseline equivalence across study groups, for individuals/couples completing pre- and mid-test, co-parenting vs. control curriculum

		COPAR	CONT	COPAR vs. CONT	Effect size
Baseline measure	n	% / Mean (SD)	% / Mean (SD)	Difference (p)	Hedge's g or Cox index
Participant Characteristics					
Female (%)	253	88%	86%	2% (0.51)	-0.11
Pregnant (vs. Parenting) (%)	253	57%	43%	14% (0.03)	0.34
Hispanic (%)	253	85%	82%	3% (0.48)	0.13
Non-Hispanic White (%)	253	10%	14%	-4% (0.28)	-0.23
Non-Hispanic Black (%)	253	10%	10%	0% (0.99)	0.00
In a relationship (%)	253	77%	75%	2% (0.73)	0.07
In a participating dyad (%)	253	24%	22%	2% (0.71)	0.07
Age	253	16.56 (1.21)	16.81 (1.18)	-0.25 (0.10)	-0.21
Birthplace = USA	253	82%	73%	9% (0.08)	0.22
Outcome Measures (Pre-Test)					
Expectations to Coparent (R 1-4)	151	3.35 (0.58)	3.42 (0.55)	0.07 (0.47)	-0.12
Coparenting Conflict (R 1-5)	189	1.89 (0.82)	2.09 (0.75)	-0.20 (0.08)	-0.25
Coparenting Interaction (R 1-5)	188	3.80 (1.18)	3.74 (1.03)	0.06 (0.71)	0.05
Healthy Relationship Attitudes A (R 1-4)	144	3.70 (0.42)	3.70 (0.33)	0.00 (0.86)	0.00
Healthy Relationship Attitudes B (R 1-4)	146	3.36 (0.56)	3.47 (0.49)	-0.11 (0.1)	-0.21
CADRI-Self (R 1-4)	181	1.53 (0.21)	1.54 (0.25)	-0.01 (0.66)	-0.04
CADRI-Partner (R 1-4)	178	1.56 (0.27)	1.62 (0.29)	-0.06 (0.19)	-0.21

Notes: COPAR = Coparenting, CONT = Control; n.a. = not applicable; p-values are include in parentheses. Effect sizes are calculated using hedge's g (continuous variables) or Cox's index (dichotomous variables). Individual baseline equivalency tests were run for each outcome variable which included all covariates AND the outcome of interest. The n next to each outcome indicates the sample size used for each individual baseline equivalence analyses. Note that "Pregnant (vs. Parenting)" is over the baseline equivalence threshold. Therefore, these samples cannot be considered equivalent and all results must be interpreted as exploratory.

Table B.3, below, shows the effect size of the baseline differences for the pre- and mid-test sample for the healthy relationship vs. control curricula. There were four differences that crossed the ES >0.25 threshold. First, students in the healthy relationships condition were more likely to be in pregnant at pre-test than students in the control condition (59% vs 43%, ES = 0.39). Second, there were fewer non-Hispanic White students in the healthy relationships curriculum (9% vs 14%, ES = 0.30). Third, more of the students in the healthy relationships curriculum were born in the US (84% vs 73%, ES = 0.27). Finally, students in the healthy relationships curriculum reported lower initial coparenting conflict than students in the control curriculum (M = 1.87 vs. 2.07, ES = 0.29). Given that the findings from the secondary research questions were

generally null, and that the research questions were secondary and not primary, we conducted the analyses as in the primary research questions section and did not rebalance the nonequivalent samples. All associated results should therefore be considered preliminary.

Table B.3. Summary statistics of key baseline measures and baseline equivalence across study groups, for individuals/couples completing pre- and mid-test, healthy relationships vs. control curriculum

		HRR	CONT	HRR vs. CONT	Effect size
Baseline measure	n	% / Mean (SD)	% / Mean (SD)	Difference (p)	Hedge's g or Cox index
Participant Characteristics					
Female (%)	246	19%	14%	0.05 (0.33)	0.22
Pregnant (vs. Parenting) (%)	246	59%	43%	0.16 (0.02)	0.39
Hispanic (%)	246	85%	82%	0.03 (0.46)	0.13
Non-Hispanic White (%)	246	9%	14%	-0.05 (0.24)	-0.30
Non-Hispanic Black (%)	246	10%	10%	0.00 (0.90)	0.00
In a relationship (%)	246	-	-	-	-
In a participating dyad (%)	246	28%	22%	0.06 (0.31)	0.19
Age	246	16.71 (1.16)	16.81 (1.18)	-0.10 (0.51)	-0.09
Birthplace = USA	246	84%	73%	-0.11 (0.04)	-0.27
Outcome Measures (Pre-Test)					
Expectations to Coparent (R 1-4)	151	3.34 (0.47)	3.42 (0.55)	-0.08 (0.35)	-0.16
Coparenting Conflict (R 1-5)	180	1.87 (0.79)	2.09 (0.75)	-0.22 (0.05)	-0.29
Coparenting Interaction (R 1-5)	180	3.92 (1.05)	3.74 (1.03)	0.18 (0.25)	0.17
Healthy Relationship Attitudes A (R 1-4)	156	3.77 (0.35)	3.70 (0.33)	0.07 (0.14)	0.21
Healthy Relationship Attitudes B (R 1-4)	156	3.50 (0.47)	3.47 (0.49)	0.03 (0.62)	0.06
CADRI-Self (R 1-4)	171	1.54 (0.22)	1.54 (0.25)	0.00 (0.91)	0.00
CADRI-Partner (R 1-4)	167	1.59 (0.33)	1.62 (0.29)	-0.03 (0.50)	-0.10

Notes: HRR = Healthy Relationships, CONT = Control; n.a. = not applicable; p-values are include in parentheses. Effect sizes are calculated using hedge's g (continuous variables) or Cox's index (dichotomous variables). Individual baseline equivalency tests were run for each outcome variable which included all covariates AND the outcome of interest. The n next to each outcome indicates the sample size used for each individual baseline equivalence analyses. Note that "Pregnant (vs. Parenting)", "Non-Hispanic White", "Birthplace", "Coparenting Conflict". Therefore these samples cannot be considered equivalent and all results must be interpreted as exploratory.

Appendix C. Survey Instruments

Coparenting attitudes

Indicate how much you agree or disagree with the statement.

	Strongly Disagree	Disagree	Agree	Strongly Agree
1. It is important for a child to have both parents in his/her life	•	•	•	•
2. It is important for a child to have multiple supportive adults in his/her life.	•	•	•	•
3. To be a good parent, sometimes you have to ask for advice from other people.	•	•	•	•
4. To be a good parent, sometimes you have to ask for help from other people.	•	•	•	•
5. It is important to have positive relationship with my child's other parent, even if we are not dating	•	•	•	•
6. It is important to share child-rearing decisions with my child's other parent	•	•	•	•
7. It is important to allow my child's other parent to spend time with our child.	•	•	•	•

Coparenting Communication & Conflict

When you and **your child's other parent** discuss parenting issues…

	Never	Rarely	Sometimes	Often	Always
1. How often does it result in an argument?	•	•	•	•	•
2. How often is there hostility and anger between you?	•	•	•	•	•
3. How often are the conversations stressful and tense?	•	•	•	•	•
4. How much do you and your child's other parent have basic differences of opinion about issues related to childrearing (parenting)?	•	•	•	•	•

How often do you talk to **your child's other parent** about the following issues:

	Never	Rarely	Sometimes	Often	Always
5. Medical issues or problems related to your baby or pregnancy.	•	•	•	•	•
6. Your baby's accomplishments and progress.	•	•	•	•	•
7. Childrearing (parenting) problems.	•	•	•	•	•
8. Planning special events for your baby.	•	•	•	•	•
9. Major decisions regarding your baby's life.	•	•	•	•	•
10. Finances in regard to your baby or pregnancy.	•	•	•	•	•
11. Daily decisions regarding your baby's life.	•	•	•	•	•

RELATIONSHIP ATTITUDES - A

In a healthy relationship, how important is it that couples…

	Not at all important	A little important	Pretty important	Very important
1. Do not cheat on each other?	•	•	•	•
2. Do not call each other names?	•	•	•	•
3. Do not threaten each other?	•	•	•	•
4. Do not push, shove, hit, slap, or grab each other?	•	•	•	•
5. Do not argue?	•	•	•	•
6. Encourage each other when life is hard?	•	•	•	•
7. Enjoy spending time together?	•	•	•	•

RELATIONSHIP ATTITUDES - B

How much do you agree or disagree with the following statements about relationships?

	Strongly Disagree	Disagree	Agree	Strongly Agree
1. In a healthy relationship it is essential for couples to talk about things that are important to them.				
2. Even in a good relationship, couples will occasionally have trouble talking about their feelings.				
3. A relationship is stronger if a couple doesn't talk about their problems.				
4. A person who makes their partner angry on purpose deserves to be hit.				
5. Sometimes physical violence, such as hitting or pushing, is the only way to express your feelings.				
6. Violence between dating partners is a personal matter and people should not interfere.				
7. It's okay to stay in a relationship even if you're afraid of your boyfriend/girlfriend.				

RELATIONSHIP CONFLICT - SELF & PARTNER REPORT

The following questions ask about things that may have happened when **you and your partner (either your current or ex-boyfriend/ex-girlfriend)** were having an argument. Please read each of the following statements and then check how many times in the **PAST MONTH <u>you</u>** have done each of the following. Do not include times when you were joking or playing around. *If you have not had a partner in the past month you may skip to page 15.*

	Never	Seldom	Sometimes	Often
1. I gave reasons for my side of the argument.	•	•	•	•
2. I tried to turn my partner's friends against him/her.	•	•	•	•
3. I did something to make my partner jealous.	•	•	•	•
4. I destroyed or threatened to destroy something my partner valued.	•	•	•	•
5. I told my partner I was partly to blame.	•	•	•	•
6. I brought up something bad my partner had done in the past.	•	•	•	•
7. I threw something at my partner.	•	•	•	•
8. I said things just to make my partner angry.	•	•	•	•
9. I gave reasons why I thought my partner was wrong.	•	•	•	•
10. I agreed that my partner was partly right.	•	•	•	•
11. I spoke to my partner in a hostile or mean tone of voice.	•	•	•	•
12. I offered a solution that I thought would make us both happy.	•	•	•	•
13. I put off talking until we calmed down.	•	•	•	•
14. I insulted my partner with put-downs.	•	•	•	•
15. I discussed the issue calmly.	•	•	•	•
16. I said things to my partner's friends about him/her to turn them against my partner.	•	•	•	•
17. I made fun of my partner in front of others.	•	•	•	•
18. I told my partner how upset I was.	•	•	•	•
19. I kept track of who my partner was with and where he/she was.	•	•	•	•
20. I blamed my partner for the problem.	•	•	•	•
21. I kicked, hit, or punched my partner.	•	•	•	•
22. I left the room to cool down.	•	•	•	•
23. I gave in, just to avoid conflict.	•	•	•	•
24. I accused my partner of flirting with another girl/guy.	•	•	•	•
25. I tried to frighten my partner on purpose.	•	•	•	•
26. I slapped my partner or pulled his/her hair.	•	•	•	•
27. I threatened to hurt my partner.	•	•	•	•
28. I threatened to end the relationship.	•	•	•	•
29. I threatened to hit my partner or throw something at him/her.	•	•	•	•
30. I pushed, shoved, or shook my partner.	•	•	•	•
31. I spread rumors about my partner.	•	•	•	•

The following questions ask about things that may have happened when **you and your partner (either your current or ex-boyfriend/ex-girlfriend)** were having an argument. Please read each of the following statements and then check how many times in the **PAST MONTH <u>your partner</u>** has done each of the following. Do not include times when your partner was joking or playing around.

If you have not had a partner in the past month you may skip to page 15.

	Never	Seldom	Sometimes	Often
1. My partner gave reasons for his/her side of the argument.	•	•	•	•
2. My partner tried to turn my friends against me.	•	•	•	•
3. My partner did something to make me feel jealous.	•	•	•	•
4. My partner destroyed or threatened to destroy something I valued.	•	•	•	•
5. My partner told me he/she was partly to blame.	•	•	•	•
6. My partner brought up something bad I had done in the past.	•	•	•	•
7. My partner threw something at me.	•	•	•	•
8. My partner said things just to make me angry.	•	•	•	•
9. My partner gave reasons why he/she thought I was wrong.	•	•	•	•
10. My partner agreed that I was partly right.	•	•	•	•
11. My partner spoke to me in a hostile or mean tone of voice.	•	•	•	•
12. My partner offered a solution that he/she thought would make us both happy.	•	•	•	•
13. My partner put off talking until we calmed down.	•	•	•	•
14. My partner insulted me with put-downs.	•	•	•	•
15. My partner discussed the issue calmly.	•	•	•	•
16. My partner said things to my friends about me to turn them against me.	•	•	•	•
17. My partner made fun of me in front of others.	•	•	•	•
18. My partner told me how upset he/she was.	•	•	•	•
19. My partner kept track of who I was with and where I was.	•	•	•	•
20. My partner blamed me for the problem.	•	•	•	•
21. My partner kicked, hit, or punched me.	•	•	•	•
22. My partner left the room to cool down.	•	•	•	•
23. My partner gave in, just to avoid conflict.	•	•	•	•
24. My partner accused me of flirting with another girl/guy.	•	•	•	•
25. My partner tried to frighten me on purpose.	•	•	•	•
26. My partner slapped me or pulled my hair.	•	•	•	•
27. My partner threatened to hurt me.	•	•	•	•
28. My partner threatened to end the relationship.	•	•	•	•
29. My partner threatened to hit me or throw something at me.	•	•	•	•
30. My partner pushed, shoved, or shook me.	•	•	•	•
31. My partner spread rumors about me.	•	•	•	•

DEPRESSION

The following statements list feelings people sometimes have. For each one, please tell me how often you have felt this way in the last month.

	Rarely or None of the time	Some or Little of the time	Occasionally or a Moderate amount	Most of the Time
1. During the past month, I was bothered by things that usually don't bother me.	•	•	•	•
2. During the past month, I felt that I could not shake off the blues even with help from my family or friends.	•	•	•	•
3. During the past month, I had trouble keeping my mind on what I was doing.	•	•	•	•
4. During the past month, I felt depressed.	•	•	•	•
5. During the past month, I felt that everything I did was an effort.	•	•	•	•
6. During the past month, I felt hopeful about the future.	•	•	•	•
7. During the past month, I thought my life had been a failure.	•	•	•	•
8. During the past month, I felt fearful.	•	•	•	•
9. During the past month, I was happy.	•	•	•	•
10. During the past month, I talked less than usual.	•	•	•	•
11. During the past month, I felt lonely.	•	•	•	•
12. During the past month, people were unfriendly.	•	•	•	•
13. During the past month, I felt sad.	•	•	•	•
14. During the past month, I could not 'get going'.	•	•	•	•

WORRY

Please indicate how well each phrase describes how you felt in the past month.

	Not at all	---	Somewhat	---	Very Much
1. My worries overwhelm me.	•	•	•	•	•
2. Many situations make me worry	•	•	•	•	•
3. I know I shouldn't worry about things, but I can't control it.	•	•	•	•	•
4. I am always worrying about something	•	•	•	•	•
5. I know that that I have been worrying about things.	•	•	•	•	•
6. Once I start worrying, I can't stop.	•	•	•	•	•
7. I worry all the time	•	•	•	•	•

Appendix D. Data Preparation

Separate data sets for demographics/background, lesson completion, nFORM pre- and post-test, and surveys (pre-, mid-, post-, and follow-up) were kept by staff. These data sets were cleaned and merged together in R software by the local evaluator. Items were coded and scale means were created in R software using the psych package; the process for this was saved in a .Rmd file so it could be audited and modified where needed. Once all data had been scaled, scored, and coded, the data file was converted to SPSS for the final analyses.

Several items to note:

- Students were required to complete 80% of items on a scale to receive a score for that scale.
- nFORM items that pertained to the local evaluation were added to the pre- or post-test data set if those items were blank in the local survey data. For example, we utilize the Healthy Relationship Attitudes scales in our evaluation. If those scales were blank in the local evaluation data set, the completed items were copied from nFORM.
- At a biweekly staff meeting, cases were discussed that needed attention. For example, in one case, a parent wrote down her baby's name in the other biological parent field, and proceeded to answer the questions on the scales that followed (which should have been answered about the other biological parent) as if they were about her baby. Her data for these scales were flagged to be removed, and a note was added to the Master List. These cases were very rare, and, when possible, the team contacted the adolescent in question to get clarifying information or to request a re-do of certain survey items.
- Names and identifying information were removed from the final file prior to analysis; a unique ID variable (the SRSF ID) was used to identify students in all files; original data were retained in case of errors or other needs to examine/merge data in the future

Missing data. All analyses used an intent-to-treat approach, such that participants who had been randomly assigned to conditions were included in the analyses regardless of their level of participation in the intervention and subsequent withdrawal or deviation from protocol. Additionally, we utilized a complete case approach by only including participants who completed both surveys of interest (e.g., pre- and post-test survey, pre- and mid-test survey) in the analyses. We applied mean imputation to adjust for missing data at the item-level. Although imputation of dependent variables is somewhat controversial, reviews of best practices (e.g., Schafer & Graham, 2002; Johnson & Young, 2011) suggest it is necessary to impute both independent and dependent variables to avoid sample bias.

Appendix E. Impact Estimation.

In this section, we present the full MLM models for each variable:

Table E.1: Expectations to Coparent: Post-test
Full Treatment vs. Partial Treatment

Initial Group Differences in Expectations to Coparent

Group Means		Number in Analysis		Test Results	
Full	Partial	Full	Partial	t	sig
3.41	3.40	62	171	-0.10	0.92

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	3.39	0.05	0.00	***	2.69	0.71	0.00	***
Group (full vs. partial)	0.01	0.10	0.92		-0.13	0.13	0.30	
Expectations to Coparent Pre-Test					0.14	0.10	0.16	
Pregnant					-0.01	0.12	0.95	
Race: White					-0.75	0.42	0.08	
Race: Black					-0.75	0.35	0.03	*
Ethnicity: Hispanic					-0.62	0.40	0.12	
Born Outside of USA					0.03	0.13	0.79	
Dyad					-0.08	0.13	0.57	
Has Romantic Partner					0.04	0.20	0.84	
Length of Relationship					0.03	0.05	0.58	
Frequency See Child's Other Parent					0.04	0.09	0.69	
Coparenting Interaction Pre-Test					0.04	0.05	0.46	
Healthy Relationships B Pre-Test					0.14	0.10	0.17	
Number of students/schools			233 / 10			163 / 10		
School level variance	Intercept only		Main model		Covariate model			
	0.00		0.00		0.00			

Source. SR/SF Post-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.2: Coparenting Conflict: Post-test

Full Treatment vs. Partial Treatment

Initial Group Differences in Coparenting Conflict

Group Means		Number in Analysis		Test Results	
Full	Partial	Full	Partial	t	sig
1.89	2.09	87	149	1.83	0.07†

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	2.09	0.07	0.00	***	0.54	0.54	0.56	
Group (full vs. partial)	-0.20	0.11	0.07	†	-0.29	-0.29	0.02	*
Coparenting Conflict Pre-Test					0.39	0.39	0.00	***
Pregnant					-0.13	-0.13	0.31	
Race: White					-0.10	-0.10	0.85	
Race: Black					0.05	0.05	0.89	
Ethnicity: Hispanic					0.19	0.19	0.68	
Born Outside of USA					0.04	0.04	0.78	
Dyad					-0.09	-0.09	0.52	
Has Romantic Partner					-0.01	-0.01	0.95	
Length of Relationship					-0.01	-0.01	0.83	
Frequency See Child's Other Parent					0.09	0.09	0.37	
Coparenting Interaction Pre-Test					-0.01	-0.01	0.85	
Healthy Relationships B Pre-Test					0.06	0.06	0.61	
Number of students/schools			236/10			196/10		
School level variance	Intercept only		Main model		Covariate model			
	0.00		0.00		0.01			

Source. SR/SF Post-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): † marginally significant at p < .10* significant at < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.3: Coparenting Interaction: Post-test

Full Treatment vs. Partial Treatment

Initial Group Differences in Coparenting Conflict

Group Means		Number in Analysis		Test Results	
Full	Partial	Full	Partial	t	sig
4.00	3.96	88	148	-0.26	0.80

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	3.95	0.10	0.00	***	0.45	1.12	0.69	
Group (full vs. partial)	0.07	0.15	0.63		0.04	0.15	0.78	
Coparenting Interaction Pre-Test					0.38	0.07	0.00	***
Pregnant					0.23	0.16	0.14	
Race: White					-0.19	0.62	0.76	
Race: Black					0.15	0.46	0.75	
Ethnicity: Hispanic					-0.07	0.56	0.91	
Born Outside of USA					0.00	0.17	0.99	
Dyad					0.06	0.17	0.74	
Has Romantic Partner					0.00	0.27	0.99	
Length of Relationship					0.10	0.06	0.09	
Frequency See Child's Other Parent					0.06	0.12	0.63	
Coparenting Conflict Pre-Test					-0.04	0.10	0.67	
Healthy Relationships B Pre-Test					0.42	0.15	0.01	**
Number of students/schools			236/10			196/10		
School level variance	Intercept only		Main model		Covariate model			
	0.00		0.00		0.01			

Source. SR/SF Post-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): † marginally significant at p < .10* significant at < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.4: Healthy Relationships A: Post-test
Full Treatment vs. Partial Treatment

Initial Group Differences in Healthy Relationships A

	Group Means		Number in Analysis		Test Results	
	Full	Partial	Full	Partial	t	sig
	3.79	3.71	99	227	-1.72	0.09†

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	3.71	0.03	0.00	***	2.66	2.66	0.00	***
Group (full vs. partial)	0.07	0.05	0.12		0.06	0.06	0.26	
Healthy Relationships A Pre-Test					0.14	0.14	0.01	*
Pregnant					0.01	0.01	0.91	
Race: White					0.38	0.38	0.07	
Race: Black					0.03	0.03	0.84	
Ethnicity: Hispanic					0.22	0.22	0.25	
Born Outside of USA					0.02	0.02	0.77	
Dyad					-0.01	-0.01	0.85	
Has Romantic Partner					-0.03	-0.03	0.69	
Length of Relationship					0.02	0.02	0.28	
Coparenting Interaction Pre-Test					0.04	0.04	0.11	
Healthy Relationships B Pre-Test					0.04	0.04	0.39	
Number of students/schools			326/10		252/10			
School level variance	Intercept only		Main model		Covariate model			
	0.0		0.0		0.0			

Source. SR/SF Post-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.5: Healthy Relationships B: Post-test
Full Treatment vs. Partial Treatment

Initial Group Differences in Healthy Relationships B

	Group Means		Number in Analysis		Test Results	
	Full	Partial	Full	Partial	t	sig
	3.45	3.39	99	227	-1.04	0.30

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	3.39	0.04	0.00	***	2.41	2.41	0.00	***
Group (full vs. partial)	0.08	0.07	0.28		0.02	0.02	0.81	
Healthy Relationships B Pre-Test					0.25	0.25	0.00	***
Pregnant					0.13	0.13	0.10	
Race: White					0.03	0.03	0.92	
Race: Black					0.04	0.04	0.85	
Ethnicity: Hispanic					-0.05	-0.05	0.87	
Born Outside of USA					-0.03	-0.03	0.75	
Dyad					-0.03	-0.03	0.72	
Has Romantic Partner					-0.24	-0.24	0.04	*
Length of Relationship					0.03	0.03	0.33	
Coparenting Interaction Pre-Test					0.08	0.08	0.02	*
Healthy Relationships A Pre-Test					0.00	0.00	0.97	
Number of students/schools			326/10		252/10			
School level variance	Intercept only		Main model		Covariate model			
	0.0		0.0		0.0			

Source. SR/SF Post-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): † marginally significant at p < .10* , * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.6: CADRI: Self: Post-test
Full Treatment vs. Partial Treatment

Initial Group Differences in CADRI: Self

	Group Means		Number in Analysis		Test Results	
	Full	Partial	Full	Partial	t	sig
	1.57	1.53	84	177	-0.96	0.34

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	1.52	0.03	0.00	***	1.17	0.34	0.00	***
Group (full vs. partial)	0.05	0.04	0.22		0.05	0.05	0.25	
CADRI: Self Pre-Test					0.38	0.38	0.00	**
Pregnant					0.03	0.03	0.56	
Race: White					-0.33	-0.33	0.07	†
Race: Black					0.26	0.26	0.06	†
Ethnicity: Hispanic					-0.40	-0.40	0.02	*
Born Outside of USA					0.00	0.00	0.97	
Dyad					0.03	0.03	0.52	
Has Romantic Partner					0.11	0.11	0.49	
Length of Relationship					-0.01	-0.01	0.51	
Coparenting Interaction Pre-Test					-0.01	-0.01	0.62	
Healthy Relationships B Pre-Test					-0.06	-0.06	0.16	
CADRI – Partner					0.17	0.17	0.06	†
Number of students/schools			261/10		197/10			
School level variance	Intercept only		Main model		Covariate model			
	0.0		0.0		0.0			

Source. SR/SF Post-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): † marginally significant at p < .10* , * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.7: CADRI: Partner: Post-test
Full Treatment vs. Partial Treatment

Initial Group Differences in CADRI: Self

Group Means		Number in Analysis		Test Results	
Full	Partial	Full	Partial	t	sig
1.62	1.59	83	174	-0.6	0.55

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	1.58	0.03	0.00	***	0.66	0.39	0.09	
Group (full vs. partial)	0.04	0.05	0.41		0.07	0.05	0.18	
CADRI: Partner Pre-Test					0.50	0.10	0.00	***
Pregnant					0.03	0.05	0.52	
Race: White					-0.09	0.21	0.67	
Race: Black					0.30	0.16	0.06	†
Ethnicity: Hispanic					-0.12	0.19	0.53	
Born Outside of USA					-0.05	0.06	0.37	
Dyad					0.08	0.05	0.15	
Has Romantic Partner					0.16	0.18	0.40	
Length of Relationship					-0.01	0.02	0.50	
Coparenting Interaction Pre-Test					-0.01	0.02	0.77	
Healthy Relationships B Pre-Test					-0.01	0.05	0.79	
CADRI – Self					0.13	0.14	0.35	
Number of students/schools					257/10		195/10	
School level variance	Intercept only		Main model		Covariate model			
	0.0		0.0		0.0			

Source. SR/SF Post-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.8: Expectations to Coparent: 3-Month Follow-Up

Full Treatment vs. Partial Treatment

Initial Group Differences in Expectations to Coparent

Group Means		Number in Analysis		Test Results	
Full	Partial	Full	Partial	t	sig
3.36	3.38	98	211	0.25	0.81

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	3.38	0.04	0.00	***	0.45	0.81	0.58	
Group (full vs. partial)	-0.02	0.07	0.81		0.02	0.12	0.87	
Expectations to Coparent Pre-Test					0.36	0.10	0.00	***
Pregnant					0.13	0.11	0.24	
Race: White					0.95	0.53	0.07	
Race: Black					0.21	0.40	0.61	
Ethnicity: Hispanic					0.74	0.50	0.14	
Born Outside of USA					-0.08	0.11	0.49	
Dyad					0.03	0.12	0.80	
Has Romantic Partner					0.02	0.19	0.92	
Length of Relationship					-0.01	0.04	0.74	
Frequency See Child's Other Parent					0.13	0.08	0.09	
Coparenting Interaction Pre-Test					-0.01	0.05	0.82	
Healthy Relationships B Pre-Test					0.10	0.10	0.28	
Number of students/schools			309 / 10			145 / 10		
School level variance	Intercept only		Main model		Covariate model			
	0.00		0.00		0.00			

Source. SR/SF 3-month Follow Up Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses. Note that this model is missing substantial data from main to covariate because the "Expectations to Coparent" scale was not introduced until follow up for the first cohort. Therefore, there are many students missing data for the covariate analyses. Sensitivity analyses removing Expectations to Coparent pretest revealed a similar pattern of results (i.e., n/s for Group).

Table E.9: Coparenting Conflict: 3-Month Follow-Up

Full Treatment vs. Partial Treatment

Initial Group Differences in Coparenting Conflict

Group Means		Number in Analysis		Test Results	
Full	Partial	Full	Partial	t	sig
1.87	1.97	80	168	0.86	0.39

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	1.97	0.07	0.00	***	-0.57	0.93	0.54	
Group (full vs. partial)	-0.10	0.12	0.39		-0.20	0.13	0.12	
Coparenting Conflict Pre-Test					0.49	0.08	0.00	***
Pregnant					-0.08	0.13	0.54	
Race: White					0.87	0.64	0.18	
Race: Black					0.11	0.47	0.82	
Ethnicity: Hispanic					0.79	0.59	0.18	
Born Outside of USA					-0.08	0.13	0.54	
Dyad					0.02	0.14	0.89	
Has Romantic Partner					-0.34	0.23	0.13	
Length of Relationship					-0.03	0.04	0.54	
Frequency See Child's Other Parent					0.10	0.10	0.35	
Coparenting Interaction Pre-Test					0.02	0.06	0.76	
Healthy Relationships B Pre-Test					0.20	0.12	0.10	†
Number of students/schools			248/10			202/10		
School level variance	Intercept only		Main model		Covariate model			
	0.00		0.00		0.03			

Source. SR/SF 3-month Follow Up Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): † marginally significant at p < .10* significant at < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.10: Coparenting Interaction: 3-Month Follow-Up

Full Treatment vs. Partial Treatment

Initial Group Differences in Coparenting Interaction

Group Means		Number in Analysis		Test Results	
Full	Partial	Full	Partial	t	sig
4.20	4.16	80	168	-0.29	0.78

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	4.16	0.08	0.00	***	3.38	1.05	0.00	**
Group (full vs. partial)	0.04	0.14	0.78		0.11	0.14	0.42	
Coparenting Interaction Pre-Test					0.40	0.06	0.00	***
Pregnant					0.00	0.14	1.00	
Race: White					-0.74	0.72	0.30	
Race: Black					-0.20	0.52	0.71	
Ethnicity: Hispanic					-0.13	0.67	0.84	
Born Outside of USA					-0.05	0.15	0.74	
Dyad					0.21	0.15	0.16	
Has Romantic Partner					0.48	0.25	0.06	
Length of Relationship					-0.03	0.05	0.51	
Frequency See Child's Other Parent					-0.24	0.11	0.03	*
Coparenting Conflict Pre-Test					-0.24	0.09	0.01	**
Healthy Relationships B Pre-Test					0.28	0.13	0.04	*
Number of students/schools					236/10		202/10	
School level variance	Intercept only		Main model		Covariate model			
	0.00		0.00		0.01			

Source. SR/SF 3-month Follow Up Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): † marginally significant at p < .10* significant at < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.11: Healthy Relationships A: 3-Month Follow-Up
Full Treatment vs. Partial Treatment

Initial Group Differences in Healthy Relationships A

	Group Means		Number in Analysis		Test Results	
	Full	Partial	Full	Partial	t	sig
	3.63	3.73	99	217	1.33	0.19

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	3.73	0.04	0.00	***	1.25	0.72	0.08	†
Group (full vs. partial)	-0.10	0.07	0.15		-0.12	0.09	0.16	
Healthy Relationships A Pre-Test					0.00	0.04	0.93	
Pregnant					0.03	0.09	0.72	
Race: White					0.81	0.46	0.08	†
Race: Black					0.37	0.32	0.25	
Ethnicity: Hispanic					0.83	0.43	0.05	*
Born Outside of USA					0.01	0.09	0.89	
Dyad					0.10	0.10	0.33	
Has Romantic Partner					-0.09	0.14	0.53	
Length of Relationship					-0.01	0.03	0.85	
Coparenting Interaction Pre-Test					0.00	0.04	0.93	
Healthy Relationships B Pre-Test					0.00	0.09	1.00	
Number of students/schools			316/10		241/10			
School level variance	Intercept only		Main model		Covariate model			
	0.0		0.0		0.0			

Source. SR/SF 3-month Follow Up Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): † marginally significant at p < .10* < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.12: Healthy Relationships B: 3-Month Follow-Up
Full Treatment vs. Partial Treatment

Initial Group Differences in Healthy Relationships B

Group Means		Number in Analysis		Test Results	
Full	Partial	Full	Partial	t	sig
3.51	3.42	99	216	-1.3	0.20

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	3.42	0.05	0.00	***	1.90	0.70	0.01	**
Group (full vs. partial)	0.08	0.07	0.30		0.04	0.09	0.62	
Healthy Relationships B Pre-Test					0.25	0.08	0.00	**
Pregnant					-0.04	0.09	0.62	
Race: White					0.06	0.45	0.89	
Race: Black					0.03	0.31	0.93	
Ethnicity: Hispanic					-0.09	0.42	0.83	
Born Outside of USA					-0.11	0.09	0.25	
Dyad					0.00	0.10	0.97	
Has Romantic Partner					-0.17	0.13	0.21	
Length of Relationship					0.00	0.03	0.93	
Coparenting Interaction Pre-Test					0.05	0.04	0.26	
Healthy Relationships A Pre-Test					0.23	0.11	0.04	*
Number of students/schools			315/10		240/10			
School level variance	Intercept only		Main model		Covariate model			
	0.01		0.0		0.0			

Source. SR/SF 3-month Follow Up Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): † marginally significant at p < .10* , * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.13: CADRI: Self: 3-Month Follow-Up

Full Treatment vs. Partial Treatment

Initial Group Differences in CADRI: Self

	Group Means		Number in Analysis		Test Results	
	Full	Partial	Full	Partial	t	sig
	1.5	1.5	73	163	0.01	0.99

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	1.50	0.03	0.00	***	0.84	0.37	0.02	*
Group (full vs. partial)	0.01	0.04	0.73		-0.02	0.04	0.67	
CADRI: Self Pre-Test					0.56	0.12	0.00	***
Pregnant					0.01	0.04	0.75	
Race: White					0.24	0.21	0.27	
Race: Black					0.19	0.15	0.20	
Ethnicity: Hispanic					0.25	0.20	0.21	
Born Outside of USA					0.00	0.05	0.95	
Dyad					0.00	0.05	0.96	
Has Romantic Partner					-0.01	0.15	0.96	
Length of Relationship					0.00	0.02	0.86	
Coparenting Interaction Pre-Test					-0.02	0.02	0.44	
Healthy Relationships B Pre-Test					-0.12	0.09	0.16	
CADRI - Partner					0.84	0.37	0.02	*
Number of students/schools			236/10			173/10		
School level variance	Intercept only		Main model		Covariate model			
	0.0		0.0		0.0			

Source. SR/SF 3-month Follow Up Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): † marginally significant at p < .10* , * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.14: CADRI: Partner: 3-Month Follow-Up

Full Treatment vs. Partial Treatment

Initial Group Differences in CADRI: Self

Group Means		Number in Analysis		Test Results	
Full	Partial	Full	Partial	t	sig
1.56	1.58	73	163	0.68	0.50

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	1.59	0.03	0.00	***	0.96	0.37	0.01	*
Group (full vs. partial)	-0.01	0.04	0.74		-0.02	0.04	0.62	
CADRI: Partner Pre-Test					0.30	0.13	0.02	*
Pregnant					0.03	0.04	0.43	
Race: White					0.25	0.22	0.25	
Race: Black					0.18	0.15	0.24	
Ethnicity: Hispanic					0.32	0.20	0.11	
Born Outside of USA					-0.05	0.05	0.32	
Dyad					0.04	0.05	0.34	
Has Romantic Partner					-0.05	0.16	0.76	
Length of Relationship					-0.02	0.02	0.31	
Coparenting Interaction Pre-Test					-0.01	0.02	0.59	
Healthy Relationships B Pre-Test					0.02	0.09	0.83	
CADRI - Self					0.96	0.37	0.01	*
Number of students/schools					236/10		173/10	
School level variance	Intercept only		Main model		Covariate model			
	0.0		0.0		0.0			

Source. SR/SF 3-month Follow Up Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.15: Depression: 3-Month Follow-Up

Full Treatment vs. Partial Treatment

Initial Group Differences in Depression

Group Means		Number in Analysis		Test Results	
Full	Partial	Full	Partial	t	sig
1.64	1.63	100	214	-0.21	0.83

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	1.59	0.03	0.00	***	1.06	0.49	0.03	*
Group (full vs. partial)	-0.01	0.04	0.74		0.05	0.07	0.53	
Depression: Pre-test					0.28	0.06	0.00	***
Pregnant					-0.01	0.07	0.87	
Race: White					0.27	0.38	0.48	
Race: Black					-0.10	0.26	0.71	
Ethnicity: Hispanic					0.21	0.35	0.56	
Born Outside of USA					0.03	0.08	0.67	
Dyad					0.23	0.08	0.01	**
Has Romantic Partner					-0.12	0.11	0.28	
Length of Relationship					-0.02	0.03	0.45	
Coparenting Interaction Pre-Test					0.01	0.03	0.81	
Healthy Relationships B Pre-Test					-0.04	0.07	0.62	
Number of students/schools					236/10		173/10	
School level variance	Intercept only		Main model		Covariate model			
	0.0		0.0		0.0			

Source. SR/SF 3-month Follow Up Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.16: Anxiety / Worry: 3-Month Follow-Up
Full Treatment vs. Partial Treatment

Initial Group Differences in Worry/Anxiety

Group Means		Number in Analysis		Test Results	
Full	Partial	Full	Partial	t	sig
2.27	2.18	100	215	-0.56	0.57

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	2.20	0.09	0.00	***	2.11	0.95	0.03	*
Group (full vs. partial)	0.04	0.14	0.76		0.04	0.15	0.81	
Worry: Pre-test					0.49	0.06	0.00	***
Pregnant					0.04	0.15	0.80	
Race: White					0.05	0.76	0.95	
Race: Black					-0.39	0.53	0.46	
Ethnicity: Hispanic					-0.14	0.70	0.85	
Born Outside of USA					0.02	0.16	0.91	
Dyad					-0.10	0.16	0.56	
Has Romantic Partner					-0.28	0.23	0.23	
Length of Relationship					0.01	0.05	0.90	
Coparenting Interaction Pre-Test					0.01	0.07	0.91	
Healthy Relationships B Pre-Test					-0.22	0.14	0.12	
Number of students/schools				315/10			238/10	
School level variance	Intercept only		Main model		Covariate model			
	0.03		0.0		0.03			

Source. SR/SF 3-month Follow Up Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.17: Expectations to Coparent: Mid-test

Coparenting vs. Control

Initial Group Differences in Expectations to Coparent

Group Means		Number in Analysis		Test Results	
CoP	Control	CoP	Control	t	sig
3.49	3.46	76	70	-0.38	0.70

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	3.46	0.06	0.00	***	1.81	0.47	0.00	***
Group (CoP vs CONT)	0.03	0.08	0.70		-0.03	0.08	0.74	
Expectations to Coparent Pre-Test					0.36	0.07	0.00	***
Pregnant					0.23	0.09	0.02	*
Race: White					-0.26	0.17	0.13	
Race: Black					0.04	0.23	0.87	
Ethnicity: Hispanic					-0.24	0.20	0.25	
Born Outside of USA					-0.07	0.10	0.49	
Dyad					0.11	0.09	0.22	
Has Romantic Partner					-0.07	0.16	0.66	
Length of Relationship					0.02	0.03	0.55	
Frequency See Child's Other Parent					0.08	0.06	0.19	
Coparenting Interaction Pre-Test					0.02	0.04	0.50	
Healthy Relationships B Pre-Test					0.02	0.07	0.79	
Number of students/schools			232 / 10			124 / 10		
School level variance	Intercept only		Main model		Covariate model			
	0.00		0.00		0.01			

Source. SR/SF Mid-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.18: Coparenting Conflict: Mid-test

Coparenting vs. Control

Initial Group Differences in Coparenting Conflict

Group Means		Number in Analysis		Test Results	
CoP	Control	CoP	Control	t	sig
1.99	1.93	98	112	-0.54	0.59

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	1.93	0.08	0.00	***	0.32	0.72	0.65	
Group (CoP vs CONT)	0.06	0.11	0.59		0.21	0.11	0.07	†
Coparenting Conflict Pre-Test					0.45	0.07	0.00	***
Pregnant					-0.29	0.12	0.02	*
Race: White					-0.25	0.28	0.37	
Race: Black					-0.29	0.29	0.31	
Ethnicity: Hispanic					-0.12	0.31	0.70	
Born Outside of USA					0.03	0.13	0.80	
Dyad					-0.10	0.12	0.41	
Has Romantic Partner					-0.07	0.21	0.75	
Length of Relationship					-0.04	0.04	0.40	
Frequency See Child's Other Parent					0.12	0.09	0.18	
Coparenting Interaction Pre-Test					0.01	0.05	0.86	
Healthy Relationships B Pre-Test					0.10	0.11	0.33	
Number of students/schools			210/10			175/10		
School level variance	Intercept only		Main model		Covariate model			
	0.00		0.00		0.01			

Source. SR/SF Mid-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): † marginally significant at p < .10* significant at < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.19: Coparenting Interaction: Mid-test

Coparenting vs. Control

Initial Group Differences in Coparenting Interaction

	Group Means		Number in Analysis		Test Results	
	CoP	Control	CoP	Control	t	sig
	4.10	3.86	97	111	-1.63	0.10†

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	3.86	0.10	0.00	***	0.53	0.89	0.55	
Group (CoP vs CONT)	0.24	0.14	0.10		0.02	0.14	0.89	
Coparenting Interaction Pre-Test					0.48	0.06	0.00	***
Pregnant					0.29	0.15	0.06	
Race: White					-0.35	0.34	0.30	
Race: Black					-0.12	0.35	0.73	
Ethnicity: Hispanic					-0.30	0.39	0.44	
Born Outside of USA					0.11	0.16	0.52	
Dyad					-0.15	0.15	0.34	
Has Romantic Partner					0.36	0.26	0.17	
Length of Relationship					0.05	0.05	0.40	
Frequency See Child's Other Parent					-0.02	0.11	0.88	
Coparenting Conflict Pre-Test					-0.07	0.09	0.43	
Healthy Relationships B Pre-Test					0.44	0.13	0.00	**
Number of students/schools			210/10		175/10			
School level variance	Intercept only		Main model		Covariate model			
	0.00		0.00		0.01			

Source. SR/SF Mid-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): † marginally significant at p < .10* significant at < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.20: Healthy Relationships A: Mid-test

Healthy Relationships vs. Control

Initial Group Differences in Healthy Relationships A

	Group Means		Number in Analysis		Test Results	
	HRR	Control	HRR	Control	t	sig
	3.77	3.64	87	70	-1.46	0.15

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	3.64	0.08	0.00	***	1.91	0.86	0.03	*
Group (HRR vs CONT)	0.12	0.09	0.18		0.13	0.11	0.22	
Healthy Relationships A Pre-Test					0.24	0.18	0.19	
Pregnant					0.21	0.12	0.09	
Race: White					0.15	0.21	0.48	
Race: Black					-0.05	0.25	0.83	
Ethnicity: Hispanic					0.26	0.23	0.27	
Born Outside of USA					-0.20	0.13	0.14	
Dyad					-0.24	0.13	0.07	
Has Romantic Partner					-0.02	0.20	0.91	
Length of Relationship					0.02	0.05	0.74	
Coparenting Interaction Pre-Test					0.07	0.09	0.45	
Healthy Relationships B Pre-Test					0.14	0.07	0.07	
Number of students/schools			157/10		126/10			
School level variance	Intercept only		Main model		Covariate model			
	0.02		0.02		0.03			

Source. SR/SF Mid-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.21: Healthy Relationships B: Mid-test

Healthy Relationships vs. Control

Initial Group Differences in Healthy Relationships B

	Group Means		Number in Analysis		Test Results	
	HRR	Control	HRR	Control	t	sig
	3.50	3.63	87	70	1.52	0.13

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	3.63	0.07	0.00	***	2.16	0.69	0.00	**
Group (HRR vs CONT)	-0.11	0.09	0.22		-0.08	0.10	0.44	
Healthy Relationships B Pre-Test					0.16	0.10	0.11	
Pregnant					0.00	0.12	0.98	
Race: White					0.18	0.20	0.37	
Race: Black					-0.03	0.24	0.89	
Ethnicity: Hispanic					0.02	0.22	0.92	
Born Outside of USA					-0.12	0.12	0.33	
Dyad					-0.17	0.12	0.16	
Has Romantic Partner					-0.07	0.18	0.71	
Length of Relationship					-0.02	0.05	0.70	
Coparenting Interaction Pre-Test					0.03	0.05	0.58	
Healthy Relationships A Pre-Test					0.27	0.15	0.07	
Number of students/schools			157/10		126/10			
School level variance	Intercept only		Main model		Covariate model			
	0.0		0.01		0.04			

Source. SR/SF Mid-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): † marginally significant at p < .10* , * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table E.22: CADRI: Self: Mid-test
Healthy Relationships vs. Control

Initial Group Differences in CADRI: Self

	Group Means		Number in Analysis		Test Results	
	HRR	Control	HRR	Control	t	sig
	1.50	1.53	87	101	0.65	0.51

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	1.52	0.03	0.00	***	1.10	0.45	0.02	*
Group (HRR vs CONT)	-0.02	0.04	0.61		-0.01	0.04	0.79	
CADRI: Self Pre-Test					0.54	0.13	0.00	***
Pregnant					-0.01	0.05	0.83	
Race: White					0.01	0.10	0.95	
Race: Black					0.10	0.11	0.37	
Ethnicity: Hispanic					0.06	0.11	0.62	
Born Outside of USA					-0.03	0.05	0.54	
Dyad					-0.02	0.05	0.68	
Has Romantic Partner					NA	NA	NA	NA
Length of Relationship					0.00	0.02	0.96	
Coparenting Interaction Pre-Test					-0.03	0.02	0.19	
Healthy Relationships B Pre-Test					-0.03	0.05	0.52	
CADRI - Partner					-0.02	0.11	0.88	
Number of students/schools			261/10		197/10			
School level variance	Intercept only		Main model		Covariate model			
	0.0		0.0		0.0			

Source. SR/SF Mid-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): † marginally significant at p < .10* , * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses. Romantic partner was dropped from the analysis due to convergence issues.

Table E.23: CADRI: Partner: Mid-test

Healthy Relationships vs. Control

Initial Group Differences in CADRI: Partner

Group Means		Number in Analysis		Test Results	
HRR	Control	HRR	Control	t	sig
1.59	1.57	86	99	-0.44	0.66

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	1.57	0.03	0.00	***	0.85	0.47	0.07	
Group (HRR vs CONT)	0.02	0.05	0.66		0.01	0.05	0.77	
CADRI: Partner Pre-Test					0.43	0.12	0.00	***
Pregnant					0.02	0.06	0.69	
Race: White					-0.06	0.12	0.62	
Race: Black					0.02	0.12	0.89	
Ethnicity: Hispanic					-0.06	0.12	0.64	
Born Outside of USA					-0.04	0.06	0.50	
Dyad					-0.02	0.05	0.68	
Has Romantic Partner					NA	NA	NA	NA
Length of Relationship					0.01	0.02	0.79	
Coparenting Interaction Pre-Test					-0.06	0.03	0.03	*
Healthy Relationships B Pre-Test					0.02	0.05	0.75	
CADRI - Self					0.12	0.14	0.41	
Number of students/schools			257/10			195/10		
School level variance	Intercept only		Main model		Covariate model			
	0.0		0.0		0.0			

Source. SR/SF Mid-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses. Romantic partner was dropped from the analysis due to convergence issues.

Appendix F. Sensitivity Analyses and Alternative Model Specifications

Table F.1 Summary statistics of key baseline measures and baseline equivalence across study groups, for individuals/couples completing pre-test and post-test, full sample

	Full Treatment	Partial Treatment	Full Treatment versus Partial Treatment	Effect size
Baseline measure	**% / Mean (SD)**	**% / Mean (SD)**	**Difference (*p*)**	**Hedge's g or Cox index**
Participant Characteristics				
Female (%)	15%	15%	0% (0.99)	0.00
Pregnant (vs. Parenting) (%)	50%	53%	-3% (0.48)	-0.07
Hispanic (%)	85%	85%	0% (0.98)	0.00
Non-Hispanic White (%)	7%	7%	0% (0.83)	0.00
Non-Hispanic Black (%)	8%	10%	-2% (0.49)	-0.15
In a relationship (%)	74%	75%	-1% (0.75)	-0.03
In a participating dyad (%)	23%	23%	0% (0.99)	0.00
Age	16.68 (1.21)	16.70 (1.16)	-2% (0.91)	-0.02
Birthplace = USA	85%	76%	9% (0.04)	0.22
Outcome Measures (Pre-Test)				
Expectations to Coparent (*R 1-4)*	3.41 (0.44)	3.31 (0.59)	0.10 (0.16)	0.18
Coparenting Conflict (*R 1-5)*	1.91 (0.79)	2.02 (0.81)	-0.11 (0.25)	-0.14
Coparenting Interaction *(R 1-5)*	3.95 (1.05)	3.72 (1.09)	0.23 (0.07)	0.21
Healthy Relationship Attitudes A (*R 1-4)*	3.73 (0.33)	3.71 (0.40)	0.02 (0.56)	<0.05
Healthy Relationship Attitudes B (*R 1-4)*	3.41 (0.51)	3.45 (0.50)	-0.04 (0.51)	-0.08
CADRI-Self (*R 1-4)*	1.53 (0.23)	1.54 (0.22)	-0.01 (0.82)	-0.04
CADRI-Partner (*R 1-4)*	1.59 (0.34)	1.59 (0.27)	0.00 (0.84)	0.00

Notes: n.a. = not applicable; *p*-values are included in parentheses. Effect sizes are calculated using hedge's g (continuous variables) or Cox's index (dichotomous variables). Full treatment N = 137, Partial = 266.

Table F.2: Expectations to Coparent: Post-test
Full Treatment vs. Partial Treatment

Initial Group Differences in Expectations to Coparent

Group Means		Number in Analysis		Test Results	
Full	Partial	Full	Partial	t	sig
3.44	3.38	89	193	-0.75	0.45

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	3.38	0.05	0.00	***	2.38	0.39	0.00	***
Group (full vs. partial)	0.06	0.08	0.47		-0.09	0.11	0.38	
Expectations to Coparent Pre-Test					0.17	0.08	0.04	*
Pregnant					0.03	0.10	0.74	
Race: Black					-0.42	0.21	0.05	*
Born Outside of USA					0.02	0.12	0.87	
Frequency See Child's Other Parent					0.04	0.02	0.09	†
Coparenting Interaction Pre-Test					0.05	0.04	0.23	
Coparenting Conflict Pre-Test					-0.10	0.06	0.09	†
Healthy Relationships B Pre-Test					0.10	0.08	0.23	
Affected by COVID					0.05	0.13	0.67	
Number of students/schools			282/ 10			196/ 10		
School level variance	Intercept only		Main model		Covariate model			
	0.00		0.00		0.00			

Source. SR/SF Post-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table F.3: Coparenting Conflict: Post-test
Full Treatment vs. Partial Treatment

Initial Group Differences in Coparenting Conflict

Group Means		Number in Analysis		Test Results	
Full	Partial	Full	Partial	t	sig
1.86	2.06	109	168	1.95	0.05 *

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	2.06	0.06	0.00	***	1.57	1.57	0.00	***
Group (full vs. partial)	-0.19	0.10	0.05	†	-0.24	-0.24	0.02	*
Coparenting Conflict Pre-Test					0.34	0.34	0.00	***
Pregnant					-0.19	-0.19	0.07	†
Race: Black					-0.10	-0.10	0.67	
Born Outside of USA					0.05	0.05	0.72	
Frequency See Child's Other Parent					-0.06	-0.06	0.04	*
Coparenting Interaction Pre-Test					0.02	0.02	0.70	
Healthy Relationships B Pre-Test					0.04	0.04	0.70	
Affected by COVID					-0.07	-0.07	0.65	
Number of students/schools			277/10			231/10		
School level variance	Intercept only		Main model		Covariate model			
	0.00		0.00		0.01			

Source. SR/SF Post-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): † marginally significant at p < .10* significant at < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table F.4: Coparenting Interaction: Post-test
Full Treatment vs. Partial Treatment

Initial Group Differences in Coparenting Interaction

	Group Means		Number in Analysis		Test Results	
	Full	Partial	Full	Partial	t	sig
	4.06	3.96	109	167	-0.84	0.40

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	3.95	0.09	0.00	***	1.05	0.60	0.08	†
Group (full vs. partial)	0.13	0.13	0.31		0.08	0.14	0.55	
Coparenting Interaction Pre-Test					0.37	0.07	0.00	***
Pregnant					0.21	0.14	0.12	
Race: Black					0.00	0.31	0.99	
Born Outside of USA					0.03	0.17	0.88	
Frequency See Child's Other Parent					0.04	0.04	0.35	
Coparenting Interaction Pre-Test					0.37	0.07	0.00	***
Healthy Relationships B Pre-Test					0.35	0.13	0.01	**
Affected by COVID					0.12	0.19	0.54	
Number of students/schools			276/10			229/10		
School level variance	Intercept only		Main model		Covariate model			
	0.00		0.01		0.04			

Source. SR/SF Post-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): † marginally significant at p < .10* significant at < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table F.5: Healthy Relationships A: Post-test
Full Treatment vs. Partial Treatment

Initial Group Differences in Healthy Relationships A

	Group Means		Number in Analysis		Test Results	
	Full	Partial	Full	Partial	t	sig
	3.81	3.72	127	250	-2.63	0.01**

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	3.72	0.02	0.00	***	2.80	0.24	< 0.00	
Group (full vs. partial)	0.10	0.04	0.02	*	0.08	0.05	0.08	
Healthy Relationships A Pre-Test					0.19	0.06	0.00	**
Pregnant					-0.02	0.04	0.70	
Race: Black					-0.06	0.10	0.53	
Born Outside of USA					-0.02	0.05	0.74	
Has Partner at Pre-Test					0.00	0.06	0.95	
Coparenting Conflict Pre-Test					0.04	0.02	0.05	
Coparenting Interaction Pre-Test					-0.01	0.03	0.80	
Healthy Relationships B Pre-Test					0.04	0.04	0.30	
Affected by COVID					0.07	0.06	0.30	
Number of students/schools			377/10		284/10			
School level variance	Intercept only		Main model		Covariate model			
	0.0		0.0		0.0			

Source. SR/SF Post-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table F.6: Healthy Relationships B: Post-test
Full Treatment vs. Partial Treatment

Initial Group Differences in Healthy Relationships B

	Group Means		Number in Analysis		Test Results	
	Full	Partial	Full	Partial	t	sig
	3.50	3.41	127	250	-1.59	0.11

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	3.42	0.04	0.00	***	2.80	0.24	< 0.00	***
Group (full vs. partial)	0.09	0.06	0.14		0.08	0.05	0.08	
Healthy Relationships B Pre-Test					0.19	0.06	0.00	***
Pregnant					-0.02	0.04	0.70	
Race: Black					-0.06	0.10	0.53	
Born Outside of USA					-0.02	0.05	0.74	
Has Partner at Pre-Test					0.00	0.06	0.95	
Coparenting Conflict Pre-Test					0.04	0.02	0.05	**
Coparenting Interaction Pre-Test					-0.01	0.03	0.80	
Healthy Relationships A Pre-Test					0.04	0.04	0.30	
Affected by COVID					0.07	0.06	0.30	
Number of students/schools			377/10		284/10			
School level variance	Intercept only		Main model		Covariate model			
	0.0		0.0		0.01			

Source. SR/SF Post-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): † marginally significant at p < .10* , * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table F.7: CADRI: Self: Post-test
Full Treatment vs. Partial Treatment

Initial Group Differences in CADRI: Self

	Group Means		Number in Analysis		Test Results	
	Full	Partial	Full	Partial	t	sig
	1.54	1.53	107	192	-0.49	0.62

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	1.51	0.03	0.00	***	2.80	0.24	< 0.00	**
Group (full vs. partial)	0.03	0.04	0.43		0.08	0.05	0.08	
CADRI: Self Pre-Test					0.19	0.06	0.00	**
Pregnant					-0.02	0.04	0.70	
Race: Black					-0.06	0.10	0.53	***
Born Outside of USA					-0.02	0.05	0.74	
Has Partner at Pre-Test					0.00	0.06	0.95	
Coparenting Conflict Pre-Test					0.04	0.02	0.05	
Coparenting Interaction Pre-Test					-0.01	0.03	0.80	
Healthy Relationships B Pre-Test					0.04	0.04	0.30	
CADRI: Partner Pre-Test					0.07	0.06	0.30	
Affected by COVID					2.80	0.24	< 0.00	
Number of students/schools			299/10		225/10			
School level variance	Intercept only		Main model		Covariate model			
	0.0		0.0		0.0			

Source. SR/SF Post-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): † marginally significant at p < .10* , * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.

Table F.8: CADRI: Partner: Post-test
Full Treatment vs. Partial Treatment

Initial Group Differences in CADRI: Partner

	Group Means		Number in Analysis		Test Results	
	Full	Partial	Full	Partial	t	sig
	1.60	1.59	106	189	-0.24	0.81

MLM Regression Models

	Main Model				Covariate Model			
Variable	B	SE	p		B	SE	p	
Intercept	1.58	0.03	0.00	***	0.55	0.32	0.09	
Group (full vs. partial)	0.01	0.04	0.72		0.05	0.04	0.28	
CADRI: Partner Pre-Test					0.47	0.10	0.00	***
Pregnant					0.04	0.04	0.32	
Race: Black					0.36	0.12	0.00	**
Born Outside of USA					-0.06	0.05	0.29	
Has Partner at Pre-Test					0.16	0.18	0.36	
Coparenting Conflict Pre-Test					-0.01	0.02	0.80	
Coparenting Interaction Pre-Test					0.03	0.03	0.42	
Healthy Relationships B Pre-Test					0.00	0.04	0.94	
CADRI: Self Pre-Test					0.08	0.13	0.54	
Affected by COVID					-0.06	0.06	0.30	
Number of students/schools			295/10			223/10		
School level variance	Intercept only		Main model		Covariate model			
	0.0		0.0		0.0			

Source. SR/SF Post-Test Survey, 2017-2019

Notes. The reference categories in the model are: partial treatment, parenting, , race: all other races, born in USA, non-dyad, no partner. Asterisks indicate the level of statistical significance ("sig"): * < 5%, ** < 1%, *** < 0.1%; NA indicates not applicable. Students must have taken both the pre- and post-test surveys to be included in the analyses.