Findings from the Replication of an Evidence-Based Teen Pregnancy Prevention Program

Evaluation of Safer Sex Intervention in New Orleans, LA

Final Impact Report for

Louisiana Public Health Institute

January 22, 2016

Prepared by

The Policy & Research Group
Eric Jenner, Ph.D.
Sarah Walsh, Ph.D.
Lynne W. Jenner, M.A.
Hilary Demby, M.P.H.
Alethia Gregory, L.M.S.W.

Recommended Citation: The Policy & Research Group (2015). *Evaluation of Safer Sex Intervention in New Orleans, LA: Findings from the Replication of an Evidence-Based Teen Pregnancy Prevention Program.* New Orleans, LA: Jenner, E., Walsh, S., Jenner, L. W., Demby, H., & Gregory, A.

Acknowledgements:

PRG would like to acknowledge the work of the LPHI TPP staff in collaborating with PRG to meet the stringent study requirements and overseeing the implementation of Safer Sex Intervention and Female Sexual Health. This effort was led by the LPHI project director, Dr. Marsha Broussard, along with Jenny Dickherber and Shantice Atkins. The health clinics that hosted the program were also essential partners, working with LPHI and PRG to integrate the study into each clinic setting. The Institute of Women and Ethnic Studies (IWES) and LPHI staff conducted fidelity monitoring. And, finally, a cadre of Health Educators, trained by IWES and LPHI staff, implemented Safer Sex Intervention and Female Sexual Health.

This publication was prepared under Grant Number 5 TP1AH000003-02-00 from the Office of Adolescent Health, 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 policies of HHS or the Office of Adolescent Health.

EVALUATION OF SAFER SEX INTERVENTION IN NEW ORLEANS, LA: FINDINGS FROM THE REPLICATION OF AN EVIDENCE-BASED TEEN PREGNANCY PREVENTION PROGRAM

I. Introduction

A. Introduction and study overview

Teens (ages 15-19) and young adults (ages 20-24) account for roughly half of all newly reported cases of sexually transmitted infections (STIs) each year in the United States, with teens at especially high risk of acquiring certain infections such as chlamydia and gonorrhea.[i] [ii] Despite precipitous declines in teen birth rates, U.S. teenagers are still at greater risk of giving birth than teens in most other developed countries.[iii] [iv] Due to the potentially deleterious social, economic, and personal outcomes associated with unintended pregnancy and the transmission of STIs (including HIV), the development and evaluation of interventions designed to reduce adolescent sexual risk through the promotion of preventative and safe sex behaviors (such as consistent condom use and abstinence) has been a priority among many researchers, health practitioners, and policy makers over the past 20 years.[ii] [iii] [v] [vi] [vii]

In 2009, the U.S. Department of Health and Human Services (HHS) began a systematic review of the evidence on programs designed to reduce and prevent teen pregnancy and STIs.[viii] Although over 35 programs have been identified as effective at reducing sexual risk behaviors (e.g., frequency of sexual activity, number of sexual partners, use of contraception) or preventing teen pregnancies and STIs, the review also highlighted the fact that few interventions have evidence of positive effects across multiple studies.[ix] In light of this, HHS incorporated replication studies into the Office of Adolescent Health (OAH) *Teen Pregnancy Prevention Program* (TPP) funding as a means to bolster the current evidence base.[8] Under the TPP grant program, the aim of replication is not simply to duplicate prior research, but rather to implement a promising program in a variety of settings, with different populations, and to evaluate whether or not the program reduces sexual risk behaviors, STIs, and teen pregnancy.[x]

One of the programs identified by the HHS TPP review as having some evidence of effectiveness is a clinic-based individualized sexual education intervention called Safer Sex Intervention. The program

was designed to reduce high-risk sexual behaviors, increase condom and contraceptive use, and prevent unplanned pregnancy and the recurrence of an STI among sexually active young women by using motivational interviewing and skills training.[xi] [xii] As evidence, the review cites one study published in 2001, that was conducted with 123 females under the age of 24 who were being treated for an STI at a pediatric hospital or an affiliated clinic in an urban area.[xi] The study was a randomized control trial (RCT) in which participants were randomly assigned to either Safer Sex Intervention or a "standard education" control condition provided at the discretion of the health provider which consisted of information related to STI transmission and condom use. Researchers conducting the study assessed a number of behavioral outcomes (sex with a main partner in the previous six months, sex with a non-main partner in previous six months, condom use at last sex, frequency and consistency of condom use with main partner, and frequency and consistency of condom use with non-main partner) via a self-report questionnaire administered at 1, 3, 6, and 12 months following baseline.

The Safer Sex Intervention study was categorized by the HHS TPP review as a "moderate-quality study with short term impact."[xiii] The review reports that there were high levels of attrition at follow-up, and only two outcomes (sex with a non-main partner and condom use at last sex) measured at one and six months after the initial session met evidence standards and were considered for the review; findings at 3- and 12-month follow-up were not considered due to sample imbalance, and findings for four measures of condom use were not considered due to the use of subgroup analyses. Results of the review indicate that the intervention had a positive statistically significant impact on having sex with a non-main partner six months after the initial session; however, no other behavioral impacts were evident.[xiii] [xiv] [xv]

In 2010, the Louisiana Public Health Institute (LPHI) received a five-year *TPP Replication of Evidence-Based Programs (Tier 1)* grant to replicate and rigorously evaluate the effectiveness of an evidence-based TPP program in New Orleans, Louisiana. LPHI contracted with The Policy & Research Group (PRG), an independent research firm, to conduct the evaluation. The objective of this study was not a direct replication of the 2001 study in which identical methods are employed to determine if the program impacted the same particular outcomes selected by the original researchers (who were more

narrowly focused on condom use and STI outcomes).[xi] Instead, the intent of the replication was more broadly focused on the substantive issue of whether or not the program (implemented as intended) was effective at reducing behaviors that put adolescents at risk for unintended pregnancy, as well as STIs.[xii] We investigate impacts on targeted behavioral outcomes (condom use, contraceptive use, frequency of sex) because they are more etiologically proximate to the intervention than the targeted health outcomes (pregnancy or STI). We reasoned that we would be more able to observe the hypothesized change in these outcomes during the time-frame of our investigation.

B. Primary research question

What is the impact of the offer to participate in Safer Sex Intervention (treatment) relative to the offer to participate in Female Sexual Health (control) on participants' reported inconsistent use of condoms six months after the end of treatment (twelve months following baseline)?

C. Secondary research questions

What is the impact of the offer to participate in Safer Sex Intervention (treatment) relative to the offer to participate in Female Sexual Health (control) on participants' (1) reported inconsistent use of contraceptives six months after the end of the intervention; and, (2) reported frequency of sex six months after the end of the intervention?

II. Program and comparison programming

The treatment condition is Safer Sex Intervention, an individual-level, clinic-based, motivational, skill-building intervention designed to reduce sexual risk behaviors, increase condom use, and prevent recurrence of STIs among adolescent females diagnosed with an STI. The control (counterfactual) condition, Female Sexual Health, is an individual-level, information-only sex education intervention that aims to increase participants' knowledge about STIs. For the purposes of this study, the two interventions (both the treatment and control conditions) were jointly referred to as the *Staying Mature and Responsible Toward Sex* (SMARTS) program.

Over the course of four years (2012-2015), LPHI collaborated with five health clinics in the New Orleans area to implement the SMARTS program. To be considered for the study, clinics had to serve the target population (females, ages 14-19), and they had to have the space and staffing to conduct the interventions. As specified in PRG's Institutional Review Board (IRB)-approved study protocols, specific names of participating study clinic sites will not be identified in this report; instead, clinics will be identified using code names of Clinic A – E.

In coordination with LPHI, each of the participating clinics hired female health educators to recruit and enroll participants, implement both Safer Sex Intervention and Female Sexual Health, and collect administrative, outcome, and fidelity-monitoring data. All health educators were expected to have either a master's degree in a science or health-related field (like public health) or a bachelor's degree in a science or health-related field with three years of relevant experience; relevant experience working with data collection, study coordination, and IRBs; and two years serving as a health educator, preferably in the field of sexual health education. Health educators were trained to implement both the Safer Sex Intervention and Female Sexual Health; they were also trained in motivational interviewing, fidelity monitoring procedures, and SMARTS research protocols. LPHI partnered with the Institute for Women and Ethnic Studies (IWES) to train health educators in intervention curricula and fidelity monitoring procedures, as well as conduct fidelity monitoring for both interventions during Years 1, 2, and 3. In Year 4, LPHI tasked former health educators to monitor fidelity.

A. Description of program as intended

The Safer Sex Intervention is based in Social Cognitive Theory, the Transtheoretical Model of Behavior Change, and motivational interviewing.[xi] The developers contend that motivational interviewing techniques can be used in clinic-based interviews to assess participants' personal sense of risk and priorities and to provide tailored health messages to participants that can more effectively promote the reduction of sexual risk behavior and to maintain that change. Coupled with skill-building exercises, the intervention is intended to increase knowledge related to risk and safe-sex behaviors, to increase

awareness of risk and need for behavior change, to help build self-efficacy to engage in safe-sex behaviors, and, ultimately, to motivate participants to engage in and maintain safe sex practices. While program content primarily focuses on condom use as an effective method to prevent both STIs and pregnancy, information is provided on other pregnancy prevention methods (e.g., oral contraceptives, spermicide), and abstinence is presented as the only 100% effective away to prevent pregnancy and the transmission of disease.

The intervention is meant to be delivered in four one-on-one sessions (an initial session and three booster sessions) over the course of six months. Each session is to be conducted in a private setting by a female health educator trained in motivational interviewing and the Safer Sex Intervention. Though the program developer indicates that intervention delivery is ideal following an STI diagnosis as it provides a "teachable moment" during which participants are more likely to be receptive to sexual health promotion messages, she also acknowledges that "the intervention can be effective if delivered to women without an STD diagnosis."[xvi]

During the initial session, participants use the "Wheel of Change" (a handout that presents personalized assessments of perceived sexual risks and safe sex behaviors), to indicate what stage of behavior change they think they are in. Health educators then lead participants through customized intervention modules based on whether they fall into the "precontemplation" or "contemplation" stage of change. A participant in the precontemplation stage of change does not see herself at risk and is not thinking about the importance of engaging in safe sex practices. By contrast, a person in the contemplation stage has considered the personal consequences of unprotected sex and has started considering how and why she should engage in safe sex behavior. Both modules cover the same 11 topics; the primary difference between the two is how the safe sex and prevention messages are delivered. The precontemplation session focuses on providing information and raising a participant's awareness of personal risk. The contemplation session focuses more on engaging in safe sex behaviors and building safe sex skills; it includes an additional role-playing activity that is not offered to those in the pre-contemplation stage.

Booster sessions delivered one, three, and six months following the initial session are intended to sustain any resulting behavior change. As with the initial session, health educators customize the booster sessions based on the participants' personal assessments of where they are on the "Wheel of Change." During booster sessions, health educators should provide an introduction, an opportunity to ask questions, a determination of the stage of change, a questions and answer period, a role play (optional), and a wrap up.

Because the intervention is to be customized to each participant, the duration of sessions may vary; typically the initial or primary intervention session should take approximately 30 to 50 minutes. Subsequent booster sessions should take 10 to 30 minutes. A more detailed description of intended program content, including number of planned activities for each session, is presented in Table A.1 in Appendix A.

B. Description of counterfactual condition

Female Sexual Health is a one-on-one, information-only sexual health intervention designed specifically for this study. It consists of a PowerPoint presentation (comprising 35 slides) that provides information on female and male reproductive anatomy, how and when a woman can become pregnant, and facts related to the cause, prevalence, and treatment of STIs including chlamydia, gonorrhea, trichomoniasis, herpes, human papillomavirus (HPV), syphilis, and HIV/AIDs. The presentation is intended to be delivered by a health educator in one face-to-face session. After the presentation, participants receive free condoms. This is the only session for the control condition; there are no booster sessions. See Table A.2. in Appendix A for an overview of presentation content.

The control condition was designed to be an information-only contrast that is time- and attention-equivalent to the initial session of Safer Sex Intervention. Both the treatment and control conditions provide to participants equivalent information regarding pregnancy, STIs, condoms, and abstinence as the only 100% effective way to prevent pregnancy and infection; the provision of such information is standard practice at many clinics providing reproductive health services to adolescents. Both conditions also involve participants in a one-on-one consultation with a female health educator for approximately 30-

50 minutes. However, the treatment (Safer Sex Intervention) differs from the control condition (Female Sexual Health) in that it attempts to motivationally engage participants with personally relevant reasons for changing behavior and build new skills and self-efficacies to use condoms, so individuals can effectively reduce their sexual risk and remain motivated to do so.

III. Study design

A. Sample recruitment

Study recruitment occurred over a two-year period, beginning in February 2012 and concluding in May 2014. Upon seeking care at one of the five study clinics, individuals meeting basic eligibility criteria (females, ages 14-19) were referred by clinicians and clinic staff to the SMARTS health educator, who conducted a full eligibility screening. Additionally, staff from other clinics or current/past study participants could refer potential participants to a SMARTS health educator to be screened for the program.

To be eligible, during the screening process, each client had to self-report to the health educator that she: a) was between the ages 14-19; b) had been recently sexually active (defined as having sex with a male in the past three months); c) was not pregnant or trying to get pregnant; and d) had not participated in a specified list of OAH TPP programs. Each client also had to indicate that she was willing to return for scheduled study sessions, and she had to provide assent/consent to participate in the study. In addition, clinicians at each clinic were able (at their own discretion) to declare individuals ineligible if they were not physically or mentally capable of participating.

Though consent requirements varied somewhat across sites, the process was the same for members of both treatment and control groups within each study site. All eligible individuals who provided the proper consent to participate were randomized and enrolled in the study at the time they attended their first scheduled study session. In all, 752 individuals were screened for eligibility. Of these, 319 met all eligibility criteria, consented, and were randomized into the study. Of the 433 individuals screened who did not enroll, 320 were deemed ineligible and were not asked to participate, and 113 were deemed eligible but did not return to the clinic for their first scheduled session. Roughly 27% (86/319) of

participants provided consent and were enrolled on the day of screening. The other 73% were screened and returned to the clinic for programming. On average, 7 days elapsed between screening and enrollment. Table B.1 in Appendix B presents details of sample recruitment across clinics including a timetable of clinics' study participation, enrollment numbers, and descriptions of consent requirements.

B. Study design

The current study involves: 1) an individual-level RCT to assess the impact of Safer Sex Intervention on sexual behaviors and 2) a descriptive assessment of the fidelity and quality of implementation to provide context in interpreting the efficacy findings. For the impact study, a blocked randomization design was used in which individuals were randomly assigned to either the treatment or control condition within each study site by way of "randomization envelopes."

Assignment occurred after consent/assent was obtained and directly before the provision of any programming or collection of baseline data. The evaluators established the randomization procedures; study staff (health educators and research assistants) were responsible for carrying out the assignments under the direction and ongoing monitoring of the evaluators. The evaluators masked the randomized assignments from the study staff through the use of an ordered set of sealed, opaque envelopes but maintained a record of the order of these assignments so that they could verify that randomization was actualized to fidelity by the study staff.

At each clinic, when enrolling and randomizing a new participant, study staff (either the health educator or research assistant) maintained the allocation sequence by selecting the randomization envelope next in the stack (next in ascending order). She recorded the study ID number from the outside of the envelope along with the participant's information in the enrollment log. Then, while the study participant was completing the baseline questionnaire, she opened the envelope and recorded the study condition to which the participant was assigned. Following completion of the questionnaire, the health educator working at the clinic at that time administered the initial intervention session corresponding to the participant's assignment. At each site, the probability of assignment to either the treatment or control group was intended to be equal; that is p (assignment to treatment) = .5.

C. Data collection

1. Impact evaluation

To assess whether or not the program affects self-reported contraceptive use and sexual activity, we use participant-reported data gathered with the SMARTS *Program Questionnaire* at baseline and six months post treatment (twelve months following baseline), which is the time period formalized in our primary and secondary research questions. Details on these instruments and their administration are provided below. Data collection procedures were the same across the two experimental conditions. At each data collection point, the treatment and control groups were asked to complete the same questionnaire, the data collection schedule and variations in mode of administration (i.e., in person, online) were offered identically across groups, and both groups were offered the same incentives to participate in data collection ($40 gift card and one entry into a semi-annual raffle for an iPod Touch or a prize of equivalent value for each questionnaire completed). Similarly, for both study conditions, health educators or research assistants recorded administrative and intervention dosage information with the same level of detail.

PRG staff constructed the SMARTS Program Questionnaire using items and scales that were adapted from those validated in prior research. The baseline questionnaire comprises 121 items, and the six-month follow-up questionnaire comprises 113 items. Both ask participants to report on various demographic characteristics, sexual behaviors, and theoretical antecedents to those behaviors. Prior to administration, the questionnaire was field-tested with 10 health professionals (including MDs, MPHs, and PhDs) as well as 6 adolescents females (ages 14 to 15) to ensure the questions were valid, relevant, and comprehendible to youth. Though slight modifications were made to the questionnaire during the study period (questions were reordered, and four questions not essential to the study were removed), no substantive changes were made. Constructs were captured with identical measures at each administration. See Appendix C for more details on the data collection efforts including modes of administration and timing of collection windows.

2. Implementation evaluation

The implementation study examined adherence to the program model, quality of implementation, experiences of counterfactuals, and context. To assess adherence to the Safer Sex Intervention, program staff collected the following administrative data for each participant: sessions offered, sessions attended, and the duration of each session. The amount of program content delivered to youth was assessed using Safer Sex Intervention Fidelity Toolkit Forms, to collect data on the specific activities completed within each session type. Forms were intended to be completed by health educators after each session and by fidelity monitor observers for a subset of audio-recorded sessions; observer data are reported whenever available, and health educator self-report data are used otherwise. Health educator credentials, employment status, study site assignment, and training completion data were provided by program administrative records.

Quality of staff-participant interactions was assessed using six questions from the OAH-required *Program Observation Form for TPP Grantees* measuring delivery of session information, extent of participants' understanding, level of participation, and overall quality of program session. These data were collected by fidelity monitor observers for a small subset (15%) of initial sessions; no observation data were collected on quality for any treatment booster sessions.

The experiences of the counterfactual condition, including sessions offered, sessions attended, and session duration, were also assessed using participant-level administrative data. The amount of Female Sexual Health counterfactual session content delivered was assessed with fidelity monitoring self-report and observer forms. Data was collected on the provision of sexual health information on 16 topic areas, as well as if the health educator engaged in any of the following five Safer Sex Intervention components: (1) motivational interviewing, (2) assessment of participant using Wheel of Change to identify how to customize session, (3) tailoring of session based on participant's feedback and stated priorities, (4) discussing consequences of sexual risk behavior and strategies to address them, and (5) teaching condom use and negotiation skills through demonstration and role-play. As with the Safer Sex

Intervention, observer data are reported whenever available, otherwise health educator self-report data are used.

The overall context of the implementation was assessed by: (1) identifying other TPP programs being implemented within the study area that were potentially available to study participants, (2) asking participants two questions about their past-year participation in other reproductive health education and TPP programs at each questionnaire data collection point, and (3) reviewing program reports and records to identify substantial unplanned adaptations to the program model. Table C.2 in Appendix C includes detailed implementation evaluation data sources, frequency of data collection, and parties responsible for data collection for each aspect of the implementation study.

Implementation evaluation data are somewhat limited by the fidelity monitoring health educator self-report and observation data gathered. Fidelity monitoring data were collected for roughly 90% of initial sessions and 60% to 83% of booster sessions. Female Sexual Health self-reported fidelity monitoring data were collected for 90% of sessions. However, for both treatment and counterfactual sessions, most of these data are self-reported by the health educators and, thus, may not be a reliable measure of the content that was actually delivered to participants. At Clinic E, only five study participants total were enrolled and no fidelity monitor self-report or observer report data were collected for these sessions, so we have no information on what these study participants received. It should be noted that fidelity monitoring was initially the responsibility of IWES, a subcontracted partner organization. LPHI discontinued their contract with IWES at the end of grant Year 3 (August 2013), after which they tasked existing program staff (former health educators) to monitor fidelity. Health educators were instructed to complete a self-report form after each session delivered and to audio record every fifth session (plus all associated booster sessions if the participant was assigned to Safer Sex Intervention); session recordings were to be provided to fidelity monitor observers for review and ongoing quality improvement feedback.

D. Outcomes for impact analyses

In this study, we investigate three behaviors which program logic suggests Safer Sex Intervention targets as a means of achieving intended health outcomes (prevention of pregnancy and STIs); these are a)

"using condoms to prevent HIV and other STIs" b) "using contraceptives to prevent pregnancy", and c) "avoiding unprotected sex by returning to abstinence."[12] As the program principally focuses on condom use as an effective risk reduction strategy, the primary research question asks whether or not the offer to participate in Safer Sex Intervention relative to the offer to participate in Female Sexual Health affects participants' inconsistency of condom use six months after the end of the intervention (12 months following baseline). Though not as central to the program, contraceptive use and avoiding unprotected sex are also covered because they are consequential to the prevention of teen pregnancy and STIs. We investigate these outcomes with secondary research questions that ask whether or not the offer to participate in Safer Sex Intervention impacts participants' inconsistency of contraceptive use (any type of contraceptives, not just condoms) and frequency of sex six months after the end of the intervention (12 months following baseline).

Primary Research Question: Inconsistency of condom use (during any type of sex)

We operationalize inconsistency of condom use as a risk outcome – the proportion of times in the past three months a participant *does not* use condoms while engaging in any type of sex (anal, oral, vaginal) – so that we may examine the self-reported sexual behaviors of the full analytic sample of participants, regardless as to whether or not they were sexually active during the study period. Study participants who indicate that they were not sexually active were assigned a proportion of 0% for this question. See Table III.1 for details of variable construction.

Table III.1. Behavioral outcome used for impact analysis of primary research question

Outcome name	Description of outcome	Timing of measure relative to program
Inconsistency of condom use	The risk outcome is operationalized as the proportion of times in the past three months a person reports having any type of sex (anal, oral, vaginal) without using a condom. The measure is calculated from the following items: • In total, how many times have you had any type of sex in the past 3 months? • Now, think about the number of times that you had any type of sex in the past 3 months. How many of those times did you use condoms? The measure is calculated by dividing the total number of times a person reported not using a condom by the total number of times she reported having any type of sex. The resulting variable is a continuous proportion with values that range from 0 to 1, where 0 indicates that a person has not engaged in sex without a condom in the past three months, and 1 indicates that the person has engaged in sex without a condom (risk behavior) 100% of the times they had any type of sex in the past three months.	6 months after treatment ends (12 months following baseline)

Secondary Research Question 1: Inconsistency of contraceptive use (during vaginal intercourse)

To be consistent with our primary research question, we also operationalize inconsistency of contraceptive use as a risk outcome – the proportion of times in the past three months a participant *does not* use any type of contraceptive (e.g., condoms, oral contraceptives, spermicide) while engaging in vaginal intercourse. See Table III.2 for details of our operationalization.

Secondary Research Question 2: Frequency of Sex (any type of sex)

Our measure of frequency of sexual activity is a continuous variable – the self-reported number of times in the past three months a person engages in any type of sex. As with our other impact analyses, in our assessment of this secondary outcome, we consider the self-reported sexual behaviors of the full analytic sample of participants, including individuals who indicate they are not currently sexually active. See Table III.2 for details of our operationalization.

Table III.2. Behavioral outcomes used for impact analyses of secondary research questions

Outcome name	Description of outcome	Timing of measure relative to program
Inconsistency of contraceptive use	The risk outcome is operationalized as the proportion of times a person reports having vaginal sexual intercourse without using any form of birth control (including condoms). The measure is calculated from the following items: • In total, how many times have you had sexual intercourse in the past 3 months? • In the past 3 months, how many times have you had sexual intercourse without using any of these methods of birth control (options listed)? The measure is calculated by dividing the total number of times a person reported having sexual intercourse without using any contraception by the total number of times she reported having sexual intercourse. The resulting variable is a continuous proportion with values that range from 0 to 1, where 0 indicates that a person has not engaged in vaginal sexual intercourse without birth control in the past three months, and 1 indicates that the person has engaged in vaginal sexual intercourse without birth control (risk behavior) 100% of the times they had sex in the past three months.	6 months after treatment ends (12 months following baseline)
Frequency of sexual activity	The risk outcome is operationalized as the number of times in the past three months a person reports having had any type of sex (anal, oral, vaginal). The measure is taken directly from the following item • In total, how many times have you had any type of sex in the past 3 months? The variable is continuous, with values ranging from 0 to k, where 0= no sexual activity reported in past 3 months and k = number of times sex reported.	6 months after program ends

E. Study sample

The full set of participants who were randomized and offered the opportunity to participate in either Safer Sex Intervention or Female Sexual Health and who provided evaluation consent/assent constitutes the full intent-to-treat (ITT) sample. Those who were randomly assigned to receive Safer Sex Intervention are considered treatment participants; those who were not are considered control participants, regardless of actual exposures. The final ITT sample consists of 319 adolescent females. The analytic sample, which is the subset of the ITT sample for whom we have sufficient data, is 268 adolescent females. Insufficient data is defined as unit non-response at baseline or six-month follow-up (an individual did not complete at least one questionnaire used in analysis), or unreliable responses at either administration, which is treated as unit missing. By unreliable we mean that data are not considered to be honest reflections of participants' behaviors because the questionnaires were completed too quickly or

participants indicated they did not answer the questions honestly. Full details of data screening, editing, and missing data procedures are specified in Appendix D.

Of the 319 youth randomized and included in the ITT sample, 37 were excluded from the study sample because they did not complete a baseline and/or six-month follow-up questionnaire; 11 were excluded because at least one of their completed questionnaires was deemed unreliable; and 3 were excluded because we do not have a signed Health Insurance Portability and Accountability Act Release Form for them (which was required by their enrollment clinic's IRB) and, therefore, cannot use their data in the evaluation. Thus, 268 participants (133 treatment; 135 control) constitute the analytic sample for both the primary and secondary analyses; this represents 84% of the full ITT sample. Demographic data collected on study participants indicate that most identify as black (81%) or multiracial (7%), and a small percentage identify as Hispanic (2%). At baseline, participants were typically 17 or 18 years old (mean = 17.5), and they reported engaging in sexual activity on average 13 times in the prior 3 months and using condoms approximately 50% of the time. Final sample sizes and response rates for each administration of the *SMARTS Program Questionnaire* as well as an outline of the sample flow for our analytic sample are included in Table B.2 in Appendix B.

F. Baseline equivalence

We assessed baseline equivalence of the treatment and control groups in the analytic sample on pre-intervention measures of our primary and secondary outcomes (inconsistency of condom use, inconsistency of contraceptive use, and frequency of sex) and five key demographic measures (age, race, ethnicity, parental education, and family structure). We used a two-step procedure to establish balance wherein we first generate model-based estimates of the differences between groups and then examine the statistical significance of the differences. Separate models were run for each of the baseline variables. Ordinary least squares (OLS) regression models were used to estimate differences in continuous baseline measures, and linear probability models were constructed to estimate differences in dichotomous baseline measures (differences are considered significant at the α=.05 level, using a two-tailed test). Results

presented in Table III.3., indicate that differences between groups are not statistically significant (i.e., $p > .05$ in all cases). Note that these baseline equivalency statistics are based on our benchmark analytic sample for which item-missing data are imputed; when we conduct the same tests on a sample for which data are not imputed results are substantively the same (see Table D.3 in Appendix D).

Table III.3. Summary statistics of key baseline measures for youth included in the 6-months post treatment analytic sample

Baseline measure	Intervention mean or proportion(sd)	Comparison mean or proportion (sd)	Mean difference	*p*-value of difference
Age (in years)	17.42 (1.35)	17.65 (1.29)	-0.23	0.14
Race: White	0.11 (0.32)	0.08 (0.27)	0.03	0.41
Race: Black	0.78 (0.41)	0.83 (0.38)	-0.04	0.35
Race: Multiracial[a]	0.07 (0.26)	0.07 (0.25)	0.01	0.83
Race: Other[b]	0.02 (0.15)	0.02 (0.15)	0.00	0.97
Ethnicity: Hispanic	0.02 (0.15)	0.07 (0.25)	-0.04	0.08
Parental education[c]	2.31 (1.03)	2.13 (0.89)	0.17	0.13
Family structure (lives with both parents)	0.10 (0.31)	0.08 (0.27)	0.02	0.54
Frequency of sexual activity[d]	13.18 (18.03)	13.67 (18.72)	-0.49	0.83
Inconsistency of condom use[e]	0.48 (0.37)	0.51 (0.4)	-0.03	0.52
Inconsistency of contraceptive use[f]	0.46 (0.29)	0.45 (0.29)	0.02	0.66
Sample size	133	135		

Notes: Reported in the table are regression adjusted means and proportions of baseline variables; standard deviations (sd) are not adjusted. Mean difference refers to the difference between the adjusted intervention and comparison means; rounding accounts for slight discrepancies in reported differences.

[a] Multiracial refers to individuals who selected more than one race category when asked "What is your race?"

[b] Race: other refers to individuals who selected one of the following race categories when asked "What is your race?": American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander, some other race.

[c] Parental education level refers to the mean level of parents' education reported by participants (1 = less than high school; 2 = high school degree or GED; 3 = associate's, technical, vocational, or trade school degree; 4 = bachelor's degree; 5 = graduate degree).

[d] Frequency of sexual activity refers to the number of times in the past three months a person reports having any type of sex.

[e] Inconsistency of condom use refers to the proportion of times in the past three months a person reports having any type of sex (anal, oral, vaginal) without using a condom.

[f] Inconsistency of contraceptive use refers to the proportion of times in the past three months a person reports having vaginal sexual intercourse without using any form of contraceptive.

G. Methods

1. Impact evaluation

The impact study investigates whether or not offering Safer Sex Intervention impacts participants' reported inconsistency of condom use (primary research question), as well as their

inconsistency of contraceptive use and frequency of sex (secondary research questions). We do this within an ITT framework, which does not adjust or account for exposure to the treatment itself. Instead, our ITT analyses allow us to estimate the effect of the offer of the treatment (Safer Sex Intervention) relative to the offer of the control condition (Female Sexual Health). To answer both primary and secondary research questions, we use a regression-estimated approach that models outcomes as a function of the baseline measure of the outcome variable (e.g., inconsistency of condom use at baseline), baseline measures of individual-level covariates that have been shown to be correlated with adolescent sexual behavior (age, race, ethnicity, parental education, and family structure) and our blocking variable (programming site). Since assignment is randomized, a simple difference of means of the outcome variables should provide an unbiased estimate of program impact; however, we use OLS regression to statistically adjust for covariates as a means to increase the precision of our estimates and to account for blocking procedures. Safer Sex Intervention is considered to have a positive impact on the self-reported sexual behaviors of participants if, at six-month follow-up (12 months following baseline), the regression adjusted means for the outcomes in the treatment group are less than those reported for the control group and the difference between the two means is statistically significant. Statistical significance is determined at the α =.05 level, using a two-tailed test. See Appendix D for details of our analytic approach, including model specifications and covariate descriptions.

Assuming that assignment procedures are conducted with fidelity, missing data pose the greatest threat to the internal validity of an RCT within an ITT framework. If we were to employ case-wise deletion in constructing our analytic sample, item-nonresponse or invalid/inconsistent responses for variables included in our empirical models would exclude between 37 and 78% of cases, depending on the research question. The number of individuals with item missing data in each of our analytic samples is as follows: inconsistency of condom use = 101, inconsistency of contraceptive use = 209, frequency of sex = 99. As outlined in our analysis plan, prior to analyzing any outcome data, we elected to minimize this loss from our analytic samples by employing dummy variable adjustment for missing pretest and

covariate data and multiple imputation for missing outcome data.[xvii] Details on our missing data approach, as well as a table describing item missingness (Table D.1), can be found in Appendix D.

In addition to our benchmark analysis described above, we conducted five sensitivity analyses to test the robustness and validity of our analytic approach. Specifically we constructed alternative empirical models or altered data cleaning and imputation rules to examine the sensitivity of benchmark findings to the following analytic decisions: (1) the use of covariates to improve the precision of our estimates; (2) the use of imputation for missing data; (3) the use of unreliable data; (4) the inclusion of outliers; and (5) the inclusion of data from persons with conflicting screening and baseline questionnaire data that pertain to eligibility. We are interested in whether or not the results produced by alternative specifications produce different inferences than the benchmark results. If they do, we conclude that the benchmark results are sensitive to our analytic decisions. Details and results of sensitivity studies are reported in Appendix E.

2. Implementation evaluation

Implementation data were analyzed using descriptive statistics to characterize program adherence and quality, as well as the counterfactual experience. To assess adherence to the program model, we present counts of sessions offered, calculate proportions and means to quantify how much was received, and list information to specify who delivered material to youth. To measure quality of staff-participant interactions, we report the percentage of observed intervention sessions in which delivery of information, participants' understanding, level of participation, and overall session quality are rated as good/moderate or better (rated 4 or 5 on response scale) by fidelity monitors. Data collected to assess the sessions offered, received, and amount of content delivered to the counterfactual condition are also quantified using counts, proportions, and means. Finally, context is depicted by listing other TPP programs available in the study area, describing unplanned adaptations to the program model, and presenting the proportion of participants (treatment and control) who report past-year exposure to other reproductive health

education and experiences with other TPP programs at each data collection point. See Table F.1 in Appendix F for detailed methods for each implementation evaluation element.

IV. Study findings

A. Implementation study findings

The implementation study provides context for the impact findings. Our study focused on four implementation elements: (1) adherence to the Safer Sex Intervention program model, (2) quality of the implementation of Safer Sex Intervention, (3) the experiences of the counterfactual group, and (4) contextual information to explain the environment in which the study was implemented and any adaptations to the program model. Below, Table IV.1 summarizes select results from the implementation study, followed by detailed findings from each implementation element.

Table IV.1. Summary of intended and actual implementation outcomes

Intended outcome	Actual outcomes
Adherence to program model: treatment participants receive one 30- to 50-minute initial session and three10- to 30-minute booster sessions at 1, 3, and 6 months after initial session.	• 100% of sessions offered • 99% received (attended) initial session; 66% received 1-month booster; 67% received 3-month booster; and 57% received 6-month booster • Average 2.9 sessions received per participant • Average initial session duration was 50 minutes and 26% lasted longer than 50 minutes • Average booster session duration was 8-10 minutes
Adherence to program model: treatment participants who identify as pre-contemplation complete 11 intervention activities; participants who identify as contemplation complete 12 activities; participants complete 6 activities at each booster session.	• Average pre-contemplation activities completed = 9.8/11 • Average contemplation activities completed = 9.8/12 • Average booster session activities completed = 4.9/6
Counterfactual group experiences: control group participants receive one 30-50 minute session called Female Sexual Health.	• 100% of sessions offered • 100% received (attended) session • Average session duration was 41 minutes • 6 participants (3.8% of control group) erroneously received Safer Sex Intervention booster session at 1-month (average duration 20 minutes)
Counterfactual group experiences: control group session provides information-only PowerPoint presentation on 16 sexual health topic areas.	• Average number of topic areas completed = 15.6/16 • Percentage of sessions in which health educator engaged in one or more Safer Sex Intervention components = 9.3%

Adherence to Safer Sex Intervention Program Model

S*essions offered and received.* Overall, 100% of intended initial, one-, three-, and six-month booster Safer Sex Intervention sessions were offered to participants assigned to the treatment group (see Table G.1 in Appendix G). Almost all participants (99%) assigned to the treatment group attended the initial session. Overall booster session attendance was considerably lower, with 66% receiving the booster session at one-month, 67% at three-months, and just 57% at six-months; there was some site-level variation in booster attendance, with a much greater proportion of study participants at sites C and D receiving all three boosters (range 64% to 86%) than those at sites A, B and E (see Table G.2). On average, participants received three (mean = 2.9) of the four intended programming sessions (see Table G.3). Only 1% of the treatment group attended no programming, and 37% attended all four sessions (see Table G.4).

The length of each session did not adhere to the expectation for the program; the average initial session duration was 50 minutes (intended 30-50), and the average booster session ranged from 8 to 10 minutes (intended 10-30 minutes) (see Table G.5). Just 28% of treatment participants received all four program sessions within the intended time frame (60-140 minutes) (see Table G.4). The majority of participants (69%) received the initial session within the intended time frame, but the session lasted more than 50 minutes for about one quarter (26%) of them. About half (51%) received booster sessions (one, three, or six) in less than 10 minutes (see Table G.6); almost all of the other booster sessions (48%) were conducted within the 10-30 minute range.

Amount of content delivered. Though data measuring the amount of content delivered have some limitations (described earlier in the implementation evaluation data collection section) and a good proportion of the initial and booster sessions were not delivered within the intended time frames, on average, it appears that most of the intended content was delivered to treatment participants; roughly 10 of 11 pre-contemplation initial session activities, 10 of 12 contemplation initial session activities, and 5 of 6 booster session activities were completed (see Table G.7). All (100%) of initial session activities were completed for 64% of pre-contemplation sessions and just 14% of contemplation sessions; all six booster

session activities were completed for just 14% of one-month, 7% of three-month, and 6% of six-month booster sessions (see Table G.8).

Program staff. A total of 11 health educators facilitated both the treatment intervention and control session; all health educators were female, met the position education requirements, and received all four required trainings: Safer Sex Intervention, Female Sexual Health curricula, fidelity monitoring, and SMARTS research protocols trainings (see Table G.9). Although the intervention was supposed to have been delivered only by trained staff, one staff member at Clinic A who was not authorized to conduct the interventions or trained in either the Safer Sex Intervention or the Female Sexual Health control condition facilitated the initial session with a total of 13 participants enrolled at this site (seven assigned to Safer Sex Intervention and six assigned to Female Sexual Health). Turnover among health educators at the various implementation study sites was relatively high. One consequence of this was that participants assigned to the treatment group (Safer Sex Intervention) did not necessarily receive all or any booster sessions from the same health educator who conducted their initial session; though receiving all four sessions from the same health educator is not a stated expectation of the program, it may be noteworthy. Further, staff turnover could also be one possible explanation for the low attendance at booster sessions.

Quality of implementation of the Safer Sex Intervention

Overall, data on quality of staff-participant interactions during Safer Sex Intervention sessions are very limited, and both the number of observations and results vary considerably by study site. Of the 23 initial sessions observed, 30% were scored as good or very good for the *delivery of session information*; *extent of participants' understanding* was scored as moderate or good in 61% of assessed sessions; *extent of group members' participation* was scored as moderate or active for 87% of assessed sessions; and *overall quality of the program session* was scored as good or excellent for 61% of assessed sessions (see Table G.10). There were no reported observation data for any treatment booster sessions.

Experiences of the counterfactual group

Sessions offered and received. In all, 100% of Female Sexual Health sessions were offered to participants assigned to the control group. All participants (100%) assigned to the control group received a Female Sexual Health session; the average session duration was 41 minutes, within the expected range of 30-50 minutes (see Table G.11). Though control group participants were intended to only receive a single session, six control group-assigned participants (3.8%) mistakenly received a Safer Sex Intervention one-month booster session (four at Clinic A, one at Clinic D, and one at Clinic C).

Amount of content delivered. On average, 15.6 of the 16 prescribed Female Sexual Health session topics were delivered; all topics were completed in approximately 84% of sessions. Overall, one or more Safer Sex Intervention components were engaged in during 9.3% of control sessions (see Table G.12).

Context

Other TPP programming. Orleans Parish was relatively saturated with other OAH-funded TPP programs during the study period and many participants reported past-year exposure to other reproductive health education, though fewer reported exposure to specific other TPP programs available in the area. There were five other TPP grantees implementing programs within Orleans Parish, though not all were targeting the same exact populations as the SMARTS study (see Table G.13). A majority of all study participants (treatment and control) reported recent (past-year) exposure to formal reproductive health education at each data collection point (See Table G.14). Although participants were screened by study staff for prior participation in specific TPP programs before being enrolled in the study, it is notable that 13% of participants (16 Safer Sex Intervention and 17 Female Sexual Health) self-reported on the baseline questionnaire that they had participated in another TPP program (other than Safer Sex Intervention or SMARTS) in the past year (see Table G.15).

Unplanned adaptations. In February of 2013, LPHI received approval from OAH to implement an adaptation to the Safer Sex Intervention to increase the length of time during which health educators could conduct one- and three-month booster sessions. The original intervention called for one- and three-month booster sessions to be held within a two-week period after the booster due date. The approved

adaptation did not alter the Safer Sex Intervention curriculum, but allowed health educators to extend the one- and three-month booster window from two weeks to one month after the due date. This adaptation was requested due to low booster session attendance; health educators were finding it difficult to get participants to return to the clinic to complete the booster sessions within the short two-week period.

B. Impact study findings

Primary Research Question: Inconsistency of condom use (during any type of sex)

Findings indicate that Safer Sex Intervention had no significant effect on participants' inconsistency of condom use at six-month follow-up (12 months after baseline). Estimates presented in Table IV.1 demonstrate statistically insignificant differences in the proportion of times treatment and control participants report using condoms when having sex in the past three months. The adjusted means for the treatment and control group of 0.50 and 0.46, respectively, indicate that at the six-month follow-up (12-months after baseline) participants in both groups had unprotected sex (any type of sex without a condom) roughly half of the time they had sex in the previous three months, on average. The adjusted mean difference between groups (0.04) is not statistically significant (p =.642). Sensitivity analyses presented in Appendix E corroborate this finding and indicate that results are not sensitive to analytical decisions. In each of the sensitivity analyses, the mean difference in participants' inconsistency of condom use reported by treatment and control groups remains statistically insignificant.

Table IV.1. Post-intervention estimated effects using data collected 6 months post treatment to address the primary research question

Outcome measure	Intervention mean (standard deviation)	Comparison mean (standard deviation)	Mean difference (p-value of difference)
Inconsistency of condom use	0.50 (0.54)	0.46 (0.74)	0.04 (0.642)
Sample Size	133	135	268

Source: Follow-up surveys administered 6 to 12 months after the treatment (Safer Sex Intervention) ends

Notes: Inconsistency of condom use refers to the proportion of times in the past three months a person reports having any type of sex (anal, oral, vaginal) without using a condom. Reported in the table are regression adjusted means and unadjusted pooled standard deviations (calculated from the 10 individual imputations used in our multiple imputation [benchmark] analysis) of the outcome variable. Included as covariates in our analytic models are the baseline measure of our outcome variable (inconsistency of condom use at baseline), age, sex, race, ethnicity, parental education level, family structure, and programming site. See Table III.3 for a more detailed description of these measures and Appendix D for details of our analytic methods, including our missing data approach.

Secondary Research Question: Inconsistency of contraceptive use (during vaginal intercourse)

Findings also indicate that Safer Sex Intervention had no impact on participants' inconsistency of contraceptive use at six-month follow-up (12 months after baseline). Adjusted means presented in Table IV.2 indicate that at the six-month follow-up (12-months after baseline) members of the treatment group on average engaged in vaginal sexual intercourse without any form of contraceptive approximately 38% of the times they had intercourse in the prior three months; members of the control group engaged in vaginal intercourse without contraceptives 23% of the time (on average). The adjusted mean difference between the two groups (0.15) is not statistically significant (p = .254). Sensitivity analyses, presented in Appendix E, again confirm this finding.

Secondary Research Question: Frequency of Sexual Activity

As with consistency of condoms and contraceptive use, results of this RCT also indicate that Safer Sex Intervention had no impact on participants' frequency of sex at six-month follow-up (12 months after baseline). Estimates presented in Table IV.2 show that treatment group members on average reported having any type of sex 18 times in the previous three months compared to 14 times reported by the control group. The adjusted mean difference between the groups of 3.7 is not statistically significant (p = .417). Sensitivity analyses, presented in Appendix E, again confirm this finding.

Table IV.2. Post-intervention estimated effects using data collected 6 months post treatment to address the secondary research questions

Outcome measure	Intervention mean (standard deviation)	Comparison mean (standard deviation)	Mean difference (p-value of difference)
Inconsistency of contraceptive use	0.38 (0.63)	0.23 (0.74)	0.15 (0.254)
Frequency of Sex	17.40 (31.52)	13.69 (39.6)	3.72 (0.417)
Sample Size	133	135	268

Source: Follow-up surveys administered 6 to 12 months after the treatment (Safer Sex Intervention) ends

Notes: Inconsistency of contraceptive use refers to the proportion of times in the past three months a person reports having sexual intercourse without using any type of contraceptives. Frequency of sex refers to the number of times in the past three months a person reports having any type of sex (anal, oral, vaginal). Reported are regression adjusted means and unadjusted pooled standard deviations (calculated from the 10 individual imputations used in our multiple imputation [benchmark] analysis) of the outcome variables. Included as covariates in our analytic models are baseline measures of our outcome variables (inconsistency of contraceptive use and frequency of sex measured at baseline), age, sex, race, ethnicity, parental education level, family structure, and programming site. See Table III.3 for a more detailed description of these measures and Appendix D for details of our analytic methods, including our missing data approach.

V. Conclusion

Findings from this RCT indicate that Safer Sex Intervention had no observable impact on certain self-reported sexual behaviors of female adolescents who were offered the program. Six months following the end of treatment (12 months from baseline), there was no statistically significant difference between treatment and control group members with regard to their self-reported inconsistency of condom use, inconsistency of contraceptive use, or frequency of sex in the past three months. Though these results appear to be inconsistent with HHS TPP evidence review findings that indicate Safer Sex Intervention can have a positive impact on some sexual behavior, this is not a direct replication of the original study (on which the evidence review's assessment is based). While both studies examine the program's impact on aspects of sexual activity and condom use, the outcomes and measures used in the two studies are not the same. In addition, the current study assesses outcomes 12 months after the initial session (6 months following the end of the intervention) whereas outcomes for the original study – that met evidence review criteria – were assessed at one and six months after the initial session.

We offer two possible explanations as to why we found no evidence of intervention impacts on inconsistency of condom use, inconsistency of contraceptive use, or frequency of sex. First, it is possible that by broadening the eligibility criteria from females recently diagnosed with or treating an STI to sexually active females (those who self-reported having any type of sex with a male partner in the previous three months), our population was not as receptive to this type of intervention. The authors of the previous study suggest that one of the reasons the intervention was shown to be efficacious is that when females have just personally experienced a negative consequence of their own risk behaviors, they are more open to health promotion and safe sex messages and more motivated to change their behavior. It is conceivable, therefore, that the population included in our study may have lacked this personal realization of vulnerability and were not as motivated to change their behavior.

Second, while brief interventions that rely on motivational interviewing techniques have been shown rather convincingly to motivate change in some types of behavior, most notably the consumption of alcohol, results from studies of brief sexual risk behavior interventions have been mixed.[xviii] [xix] [xx] [xxi]

Meta-analyses of RCTs of motivational interviewing interventions show that a number of interventions using these methods to curb or modify risky sexual behaviors have been found to be ineffective, and, when impacts are apparent, they tend to be small.[xviii xix xx xxi] Findings of effectiveness might, in other words, be uneven across time and/or populations because the average effects tend to be small in magnitude.

In any case, study results such as this that fail to reject the null hypothesis and find that the intervention has failed to affect the behavioral change hypothesized should be of equal evidentiary value to those that find otherwise. They may, in fact, provide more opportunity or incentive to learn why the intervention works in some cases and not in others and what conditions are necessary for causal impacts. The goal of OAH-funded TPP evaluations is to build our knowledge base of what works in teen risk reduction. Though there is evidence that Safer Sex Intervention can be effective at decreasing risk behaviors, findings in this study indicate that the program was not effective in this setting, with this specific population.

This study is necessarily limited in its scope. We investigate whether or not a program identified as having evidence of effectiveness has a causal effect on select sexual behaviors that are measurable and known to increase the risk of contracting STIs and unintended pregnancy. The findings may not be immediately generalizable beyond our analytic sample (sexually active females ages 14-19 seeking care at health clinics in New Orleans), the outcomes analyzed, or the specific program implemented. The results, however, do contribute to the research base for Safer Sex Intervention and, when considered with evidence from other studies, should help researchers and programmers to understand what the common or average effect of Safer Sex Intervention may be. Moreover, when compiled and synthesized with findings from other studies, this research may bring us closer to a broader understanding of what works to reduce sexual behaviors that puts adolescents' health at risk. They, moreover, should help us better understand how these programs work, with whom, and under what conditions.

VI. References

[i] Centers for Disease Control and Prevention. Sexually Transmitted Disease Surveillance 2012. Atlanta: U.S. Department of Health and Human Services; 2013. http://www.cdc.gov/std/stats12/surv2012.pdf. Accessed June 15, 2015.

[ii] Centers for Disease Control and Prevention. HIV Surveillance Report, 2013; vol. 25.

[iii] Centers for Disease Control and Prevention. About Teen Pregnancy: Teen Pregnancy in the United States.

[iv] Martinez G, Copen CE, Abma JC. Teenagers in the United States: Sexual Activity, contraceptive use, and childbearing 2006-2010. National Survey of Family Growth. National Center for Health Statistics. Vital Health Statistics. 2011; 23(31). http://www.cdc.gov/nchs/data/series/sr_23/sr23_031.pdf. Accessed June 15, 2015.

[v] Centers for Disease Control and Prevention. CDC fact sheet: incidence, prevalence, and cost of sexually transmitted infections in the United States. http://www.cdc.gov/std/stats/STI-Estimates-Fact- Sheet-Feb-2013.pdf. Accessed June 15, 2014.

[vi] The National Campaign to Prevent Teen and Unplanned Pregnancy. Counting it up: key data. https://thenationalcampaign.org/resource/counting-it-key-data-2013. Accessed June 15, 2014.

[vii] Owusu-Edusei K, Chesson H, Gift T, et al. The estimated direct medical cost of selected sexually transmitted infections in the United States, 2008. *Sex Transm Dis*. 2013; 40(3):197-20

[viii] TPP Resource Center: Evidence-Based Programs. US Department of Health and Human Services, Office of Adolescent Health Web site. http://www.hhs.gov/ash/oah/oah-initiatives/teen_pregnancy/db/index.html. Accessed August 20, 2015.

[ix] Goesling B, Colman S, Trenholm C, et al. Programs to reduce teen pregnancy, sexually transmitted infections, and associated sexual risk behaviors: A systematic review. *J Adolesc Health*. 2012; 54:499-507.

[x] Goesling B. Making Sense of Replication Studies: Guidance for Teen Pregnancy Prevention Researchers. ASPE Research Brief, 2015. http://aspe.hhs.gov/pdf-report/making-sense-replication-studies-guidance-teen-pregnancy-prevention-researchers. Accessed June 15, 2015.

[xi] Shrier LA, Ancheta R, Goodman E, et al. Randomized controlled trial of a Safer Sex Intervention for high-risk adolescent girls. *Arch Pediatr Adolesc Med*. 2001 Jan; 155(1): 73-9.

[xii] Firpo-Triplett R, Rex P, Shrier L. Safer Sex Intervention adaptation kit. Scotts Valley, CA: ETR Associates. http://recapp.etr.org/recapp/documents/programs/SSI_Adaptation_Kit.pdf. Published 2011. Accessed June 15, 2015.

[xiii] Mathematica Policy Research. Identifying Programs That Impact Teen Pregnancy, Sexually Transmitted Infections, and Associated Sexual Risk Behaviors: Review Protocol. Version 1.0. http://tppevidencereview.aspe.hhs.gov/pdfs/PPRER_Protocol_7-27-2011.pdf. Accessed June 15, 2015.

[xiv] Research Evidence for Safer Sex Intervention. Teen Pregnancy Prevention Evidence Review Web site, 2012. http://tppevidencereview.aspe.hhs.gov/pdfs/SaferSexIntervention.pdf. Accessed August 20, 2015.

[xv] Safer Sex Study Details. Teen Pregnancy Prevention Evidence Review Study Database. http://tppevidencereview.aspe.hhs.gov/StudyDetails.aspx?id=275. Accessed August 20, 2015.

[xvi] Safer Sex Intervention Program Frequently Asked Questions. Boston Children's Hospital Web site. http://www.childrenshospital.org/centers-and-services/safer-sex-intervention-program/frequently-asked-questions. Accessed August 20, 2015.

[xvii] Puma MJ, Olsen RB, Bell SH, Price C. What to do when data are missing in group randomized controlled trials. National Center for Education Evaluation and Regional Assistance: Institute of Education Sciences. Washington, DC: US Department of Education; 2009:0049.

[xviii] Burke B, Arkowitz H, Menchola M. The efficacy of motivational interviewing: a meta-analysis of controlled clinical trials. *J Consult Clin Psychol.* 2003 Oct; 71(5):843-61.

[xix] Lundahl B, Moleni T, Burke BL, et al. Motivational interviewing in medical care settings: a systematic review and meta-analysis of randomized controlled trials. *Patient Educ Couns*. 2013 Nov; 93(2):157-68.

[xx] Dunn C, Deroo L, Rivara F P. The use of brief interventions adapted from motivational interviewing across behavioral domains: a systematic review. *Addiction*. 2001 Dec; 96(12):1725-42.

[xxi] Lundahl BW, Kunz C, Brownell C, Tollefson D, Burke BL. A meta-analysis of motivational interviewing: twenty-five years of empirical studies. *Research on Social Work Practice.* 2010; 20(2): 137-160.

Appendix A: Details of Intervention Content

Table A.1 Intended Program Content for Safer Sex Intervention, by Session

Activity	Overview
Initial Session:	
1. Introduction to Safer Sex Intervention	The health educator provides an overview of the intervention and asks the participant to watch a short video.
2. Stage of Change Determination	Using the "Wheel of Change," the health educator then guides the participant through an activity to determine her stage of change and they discuss the participants concerns related to unprotected sex.
3. Consequences of Unprotected Sex	The health educator then provides information on STIs and discusses the consequences of unprotected sex.
4. Risk Perception	The participant and health educator engage in a discussion of the participants personal risk perceptions.
5. Preventing the Consequences	The participant and health educator discuss methods to prevent pregnancy; they discuss birth control options and the ways in which these can (or cannot) prevent STIs.
6. Condoms	The participant is asked about her use of condoms; she is shown how to use both male and female condoms, and she is asked to practice applying the male condom to a model.
7. Obtaining Condoms	The participant and health educator discuss how and where to obtain condoms and personal obstacles to obtaining them.
8. Secondary Abstinence	The participant and health educator discuss secondary abstinence as the only 100% effective way to prevent pregnancy and STIs, and they discuss ways to say no to sex.
9. Talking About Sex	The participant and health educator talk about pressures to have sex and the importance of talking with your partner about sex.
10. Role Play	***Contemplation module only** The health educator engages the participant in a role playing activity in which the participant practices talking about sex with her partner.
11. Questions and Answers	At the end of the session, the participant is provided the opportunity to ask questions.
12. Feedback and Summary	The health educator provides feedback on the session and summarizes what was covered. The participant is able to take all brochures and handouts distributed during the session and is provided condoms if appropriate.
Booster Session 1, 2, & 3	
1. Introduction	The health educator begins by building/reestablishing rapport with the participant.
2. Questions to Start	The participant is then asked questions about her recent sexual practices.
3. Stage of Change Determination	The participant is asked to determine her stage of change using the "Wheel of Change." They discuss the participants concerns regarding safe sex and if there is anything she would like to improve.
4. Questions and Answers	The health educator provides the participant the opportunity to ask questions.
5. Role Play (optional)	If appropriate to the participants stage of change, the participant and health educator engage in the role-playing activity to practice talking about sex.
6. Wrap Up	At the end of the session, the participant is offered condoms and intervention materials from the first session.

Table A.2 Intended Program Content for Female Sexual Health

Topic area/activity	Overview
Introduction	The health educator provides an overview of the content of the presentation.
1. Female reproductive anatomy	The health educator goes through 1 slide that covers female reproductive anatomy; she informs the participants of how and when she can become pregnant.
2. Male reproductive anatomy	The health educator goes through 1 slide that covers male reproductive anatomy.
3. STIs	The health educator goes through 2 slides that provide an overview of STIs including their prevalence, how they can be passed, and how many there are.
4. Chlamydia	The health educator goes through 4 slides that provide detailed information on Chlamydia symptoms, prevalence, and treatment.
5. Gonorrhea	The health educator goes through 3 slides that provide detailed information on Gonorrhea symptoms, prevalence, and treatment.
6. Trichomoniasis	The health educator goes through 2 slides that provide detailed information on Trichomoniasis symptoms, prevalence, and treatment.
7. Herpes	The health educator goes through 2 slides that provide detailed information on Herpes symptoms, prevalence, and treatment.
8. Human Papillomavirus (HPV)/Cervical Cancer	The health educator goes through 3 slides that provide detailed information on HPV and cervical cancer symptoms, prevalence, and treatment.
9. Syphilis	The health educator goes through 3 slides that provide detailed information on Syphilis symptoms, prevalence, and treatment.
10. STIs overview	The health educator goes through 1 slide that provides an overview of types of STIs (bacterial, viral, fungal, parasites).
11. HIV/AIDS - introduction	The health educator goes through 2 slides that provide an introduction to HIV/AIDS.
12. HIV/AIDS - transmission	The health educator goes through 1 slide that provides detailed information on HIV/AIDS transmission.
13. HIV/AIDS – rates in teens and women	The health educator goes through 4 slides that provide statistics on HIV/AIDS rates by state and prevalence in teens and women and 1 slide on HIV/AIDS treatment.
14. Prevention	The health educator goes through 1 slide that provides an overview of prevention practices including getting tested, using condoms and lubricant, talking with a partner about sex and STIs, and abstinence as the only 100% effective method of prevention.
15. Questions	The health educator asks the participant if she has any questions about the information provided.
16. Provide participants with condoms	The health educator offers participant condoms.

Appendix B: Study sample

Prior to implementation LPHI partnered with two clinic sites (Clinics C and E) that expected to be able to recruit and enroll 610 participants into the study. However, after the study began, it became clear there were fewer eligible patients receiving services at these partner clinics than originally anticipated. LPHI attempted to remedy this by partnering with several additional clinics during the study period (LPHI subsequently partnered with Clinic A, followed by Clinic B, and finally Clinic D) to increase participant enrollment. The partnership with Clinic E was discontinued after two months because staff had enrolled only five study participants in that time. The partnership with Clinic A was discontinued after five months due to high turnover rates among health educators and poor study management. Though the participants enrolled at Clinics A and E remained in the study (i.e., study staff continued to follow-up with participants), recruitment efforts ceased at both sites. Below we present descriptive statistics on recruitment, screening, and enrollment across clinics.

Table B.1 Recruitment, Screening, and Enrollment, by Clinic

Clinic	Enrollment dates	Projected Enrollment	Number recruited	Number screened	Number ineligible	Number eligible not enrolled	Number enrolled	Consent Requirements[e]
Clinic A	6/11/2012 - 10/31/2012	100	65	65	11	7	47	Participant Assent
Clinic B	6/9/2012 - 4/1/2013	100	55	52	6	17	29	Participant Assent, Parental Consent for Participants under the age of 18, and signed HIPAA Release Form
Clinic C	3/5/2012 - 2/28/2014	305	335	325	182	50	93	Participant Assent, Parental Consent for Participants under the age of 18, and signed HIPAA Release Form
Clinic D	7/5/2013 - 5/31/2014	180	419	292	120	27	145	Participant Assent
Clinic E	2/6/2012 - 4/4/2012	305	18	18	1	12	5	Participant Assent
Total		610[a]	892	752[a]	320	113[c]	319[d]	

Notes: [a] Initial projected enrollment for the project was 610 based on enrollment expectations for the two original implementation sites, Clinic C and E. [b]140 people were not interested in being screened for the study (3 at Clinic B, 10 at Clinic C, and 127 at Clinic D). [c] Persons who were eligible but did not enroll were eligible based on the screening and were asked to participate. Though they indicated they were willing to come in for scheduled sessions, following the screening, these individuals did not provide required consent materials, and they did not return to the clinic to attend their first session. [d] Participants could be recruited through the clinic or another source. In all, nine individuals were referred to a clinic from another source – four enrolled at Clinic C, one enrolled at Clinic D, and four enrolled at Clinic E. [e] Participants ages,14-17 were required to provide Participant Assent to participate in the study; participants ages 18 or 19 were required to provide Participant Consent. Parental Consent was only required for Clinic B and Clinic C sites for participants under the age of 18; Clinic B and Clinic C's Institutional Review Board also required that a Health Insurance Portability and Accountability Act (HIPAA) Release Form be signed by participants and, if applicable, a parent (when participants were under age 18).

Table B.2. Sample Flow for Benchmark Analytic Sample, by Intervention Status

	Total sample size	Intervention sample size	Comparison sample size	Total response rate	Intervention response rate	Comparison response rate
Number of youth:						
Assigned to condition	319	159	160			
Contributed a baseline survey[a]	314	157	157	98.4%	98.7%	98.1%
Contributed a 6-month follow-up survey	281	140	141	88.1%	88.1%	88.1%
In benchmark analytic sample: Contributed reliable baseline and 6 month follow-up questionnaires (item-missing data imputed)[b]	268	133	135	84.0%	83.6%	84.4%
In analytic sample and contributed outcome responses (item missing outcome data not imputed): Inconsistency of condom use[c]	209	109	100	65.5%	68.6%	62.5%
In analytic sample and contributed outcome responses (item missing outcome data not imputed): Inconsistency of contraceptive use [c]	128	62	66	40.1%	39.0 %	41.3%
In analytic sample and contributed outcome responses (item missing outcome data not imputed): Frequency of sex [c]	211	110	101	66.1%	69.2%	63.1%

Notes: [a]Three individuals (two from Clinic B and one from Clinic C) were enrolled and assigned to a condition, but they did not sign a HIPAA Release Form, which was required by the Clinics' IRB. These individuals' outcome data cannot be used in the evaluation; therefore, we consider them in this table as not having completed the baseline questionnaire. [b]This row of data reflects our benchmark analytic sample – the group of individuals who completed reliable questionnaires and whose behaviors are reflected in impact findings. Both baseline and outcome item-missing data are imputed for this sample. [c]This row of data reflects the number of individuals in our analytic sample who provided responses for outcome measures; that is, it excludes individuals who were missing data needed to construct the outcome variable.

Appendix C: Data collection efforts

Table C.1. Participant data collection windows at baseline and 6-month follow-up

	Baseline Questionnaire	6-Month Follow-up Questionnaire
Data Collection Window Opened	On study enrollment date; administered prior to the first Safer Sex Intervention or Female Sexual Health session	6 months following the final treatment booster session (12 months following enrollment date)
Data Collection Window Closed	30 days after youth enrolled in study (30 days following enrollment date)	1 day before the 12-month follow-up questionnaire opened (18 months following enrollment date)

Participants were encouraged to complete all questionnaires (baseline and follow-up) in person at a clinic. In the clinic setting, questionnaires were administered electronically (using a computer) in a quiet, private space using a web-based (online) survey administration platform that had Audio Computer Assisted Self-Interview capabilities (i.e., participants had the option to listen to audio recordings of all questions and response options). Study staff (either the health educator or a research assistant) set up the questionnaire (i.e., she accessed the questionnaire to be completed and entered the participant's study ID into the electronic form) and gave the study participant brief instructions about how to complete it. The script for the instructions emphasized the importance of the participant's honesty in answering questions and the confidentiality of her responses. Study staff also provided the participant with a sheet of paper containing definitions for relevant terms and a calendar to reference. No study staff were present in the room while the participant completed the questionnaire. If, for some reason, the computer was not working, the study participant completed the questionnaire using a self-administered paper form otherwise identical to the electronic version. On occasion, if it was inconvenient for participants to come into the clinic, study staff would meet participants at locations other than the clinic (such as libraries or coffee shops) to complete the questionnaire in person; procedures were the same as those used when completing the questionnaire in person at the clinic sites.

If a participant was unwilling or unable to complete a follow-up questionnaire in person, she also had the option to complete it online starting one month after her data collection window opened or by phone in an interview format starting five months after the window opened. If completed online outside of the clinic, a link to the questionnaire and instructions were emailed to the participant; the participant's study ID was embedded in the online form so that the data were linkable to that participant. The online version of the questionnaire was identical to that used in the clinic. The phone interview mode of questionnaire administration used an abbreviated version of the questionnaire (28 questions). It contained those questions necessary for our impact analysis (i.e., questions from which our outcome measures are constructed) as well as select questions gauging participants' perceptions and attitudes associated with safe sex practices.

Administrative study data were collected with the *Recruitment Log* and the *Enrollment* Log. Both logs were maintained by health educators electronically using a Microsoft Access database. Each study site (clinic) had its own set of logs. If a youth was screened for eligibility, eligibility data – including a numeric recruitment number, screening date, recruitment method, self-reported age, and reasons not eligible – were entered from a paper *Recruitment Form* into the electronic *Recruitment Log* for each potential participant who was screened. Once a participant was enrolled in the study, relevant study information – including the participant's recruitment number, Study ID, attendance data, questionnaire completion information, and data collection due dates – were recorded in the *Enrollment Log*. A password-protected copy of each clinic's *Recruitment Log* and *Enrollment Log* (minus personally identifiable information) was submitted to the evaluators weekly; the evaluators maintained these data in separate databases linkable by the Study ID number.

Table C.2. Data used to address implementation research questions

Implementation element	Types of data used to assess whether or not the element of the intervention was implemented as intended	Frequency/sampling of data collection	Party responsible for data collection
Adherence to Safer Sex Intervention program model: How many sessions were offered?	The **Enrollment Log** collects the following attendance data for all four Safer Sex Intervention sessions (initial session, 1-month booster, 3-month booster, and 6-month booster): session date, session conducted (yes or no), length of session (in minutes).	Data in the Enrollment Log were collected for every session that was offered.	Enrollment Log data were recorded by Program staff (Health Educator) who offered the session; data files were maintained separately for each study site.
Adherence to Safer Sex Intervention program model: What and how much was received?	The **Enrollment Log** collects the following attendance data for all four Safer Sex Intervention sessions (initial session, 1-month booster, 3-month booster, and 6-month booster): session date, session conducted (yes or no), length of session (in minutes)	Data in the Enrollment Log were collected for every session that was offered.	Enrollment Log data were recorded by Program staff (Health Educator) who offered the session; data files were maintained separately for each study site.
Adherence to Safer Sex Intervention program model: What amount of content was delivered to youth?	The **Safer Sex Intervention Fidelity Toolkit Self-Report Form** collects the following administrative and fidelity data for each session (initial session and three booster sessions): participant ID, site name, session date, total session time; session activity completed, activity not completed, or not applicable; comments about each activity; qualitative data about adaptations to and/or issues with session.	Safer Sex Intervention Fidelity Toolkit Self-Report data were intended to be recorded after each intervention session.	Self-Report fidelity data were intended to be reported by Program staff (Health Educator) following each session that is delivered.

Implementation element	Types of data used to assess whether or not the element of the intervention was implemented as intended	Frequency/sampling of data collection	Party responsible for data collection
	The **Safer Sex Intervention Fidelity Toolkit Observer Form** collects observation data for each audio recorded intervention session (same administrative and fidelity data collected as with self-report form; data collection dependent on participant providing consent to be audio-recorded).	For new Health Educator staff and new study sites: audio-recordings were collected and Safer Sex Intervention Fidelity Toolkit Observer data were intended to be recorded for all intervention sessions and their associated booster sessions for the first month of implementation. For existing Health Educators and study sites (implementation ongoing for more than one month): audio-recordings were intended to be collected and Safer Sex Intervention Fidelity Toolkit Observer data were intended to be recorded for every fifth session plus all associated booster sessions with selected participants.	Observer fidelity data were intended to be reported by Program staff (Fidelity Monitors) following review of each recorded session that was delivered.
Adherence to Safer Sex Intervention program model: Who delivered material to youth?	List of Health Educators hired to implement program at each study site, including their credentials (degree/certifications) and employee type (intern, part-time, or full-time employee).	Health Educator employment, credentials, and employee type data were available to administrative program staff.	Health Educator data were maintained by Administrative Program staff (Project Manager and Health Education Manager).
	List of Health Educator position qualification requirements (as created by program staff).	Health Educator position requirements were determined prior to hire date and available to program staff.	Health Educator data were maintained by Administrative Program staff (Project Manager and Health Education Manager).
	Lists of Health Educator staff who completed the following trainings: Safer Sex Intervention, Female Sexual Health control condition, Fidelity Monitoring Procedures, and SMARTS Research Protocols.	Training attendance data were available to administrative program staff.	Health Educator data were maintained by Administrative Program staff (Project Manager and Health Education Manager).
	The Enrollment Log collects data on the site where each health educator implemented the program.	Health educator data in the Enrollment Log were collected for every session that was offered.	Enrollment Log data were recorded by Program staff (Health Educator) who offered the session; data files were maintained separately for each study site.

Implementation element	Types of data used to assess whether or not the element of the intervention was implemented as intended	Frequency/sampling of data collection	Party responsible for data collection
Quality: Quality of staff-participant interactions	The **Program Observation Form for TPP Grantees** (developed by OAH) collects data to assess the overall quality of the program session and delivery of the information.	Program Observation Form data were recorded for each audio-recorded initial Safer Sex Intervention session (not recorded for booster sessions).	Program Observation Form data were reported by Fidelity Monitors following review of each recorded session that was delivered.
Counterfactual comparison condition experiences: What was offered and received?	The **Enrollment Log** collects the following attendance data for Female Sexual Health: session date, session conducted (yes or no), length of session (in minutes).	Data in the Enrollment Log were collected for every session that was offered.	Enrollment Log data were recorded by Program staff (Health Educator) who offered the session; data files were maintained separately for each study site.
Counterfactual comparison condition experiences: What amount of content was delivered to youth?	The **Female Sexual Health Fidelity Monitoring Self-Report Form** collects the following self-reported administrative data and fidelity data for the counterfactual session: participant ID, site name, session date, total session time; facilitator provided information on each topic (yes or no); qualitative comments about each topic; facilitator engaged in any assessed components of Safer Sex Intervention (yes or no; comments); facilitator comments about the session.	Female Sexual Health Fidelity Monitoring Form data were intended to be recorded after each session.	Self-Report fidelity data were intended to be reported by Program staff (Health Educator) following each session that was delivered.
	The **Female Sexual Health Fidelity Monitoring Observer Form** collects observation data for each audio recorded session (same administrative and fidelity data collected as with self-report form; data collection dependent on participant consent to be audio-recorded).	For new Health Educator staff and new clinic sites: audio-recordings were intended to be collected and Fidelity Observer data were to be recorded for all sessions for the first month of implementation. For existing Health Educators and sites: audio-recordings were intended to be collected and Fidelity Observer data were to be recorded for every fifth session.	Observer fidelity data were reported by Program staff (Fidelity Monitors) following review of each recorded session that was delivered.
Counterfactual comparison condition experiences: Who delivered material to youth?	Same as listed in adherence section.	Same as listed in adherence section.	Same as listed in adherence section.

Implementation element	Types of data used to assess whether or not the element of the intervention was implemented as intended	Frequency/sampling of data collection	Party responsible for data collection
Context: Other TPP programming available or offered to study participants (both intervention and comparison)	List of other TPP programming being implemented in Orleans Parish during program period.	List initially developed during grant year one and updated on an ongoing basis if new programs were identified.	List of other TPP programs is recorded by Evaluation staff (PRG Research Analyst).
	Two items on the **SMARTS Questionnaire** collect individual-level self-reported data on participants' reproductive health education and experiences with other TPP programs in the past year.	Questionnaire data were collected from participants at baseline and 6-months following the end of treatment (12months postbaseline.)	Questionnaire data were collected by both Program staff (Health Educators) and Evaluation staff (PRG Research Assistant).
Context: External events affecting implementation	News sources indicating external events that affect implementation.	Ad hoc	News sources data were to be recorded by Evaluation staff (PRG Research Analyst and Research Assistant).
	The **SMARTS Study Methods Log** collects data about external events that may affect implementation.	Ad hoc	Methods Log data were recorded by Evaluation staff (PRG Research Analyst and Research Assistant).
Context: Substantial unplanned adaptations	Adaptation requests to OAH, OAH progress reports (6-month and annual), SMARTS project meeting notes.	Adaptation requests were completed as needed; progress reports completed every six months; meeting notes taken at bi-weekly project meetings.	Adaptation requests, progress reports, and some meeting notes were recorded by Administrative staff (Project Manager and Health Education Manager). Some meeting notes were recorded by Evaluation staff (independent evaluator PRG).

TPP = Teen Pregnancy Prevention.

PRG = The Policy & Research Group

Appendix D: Methods

Model Specification

The empirical models for each research question were estimated with an OLS regression using Stata (StataCorp. Statistical Software: Release 13. College Station, TX: StataCorp LP). We present the empirical model for our primary research question below; the models for our secondary research questions are identical except that the outcome variable and baseline measure of the outcome variable is inconsistency of contraceptive use (continuous proportion; range 0 to 1) or frequency of sex (continuous, range 0 to *k*).

$$Y_{Post} = \beta_0 + \beta_1 T + \beta_2 Y_{Pre} + \sum \beta_p X_p + \varepsilon$$

where:

Y_{Post} – The outcome variable, inconsistency of condom use (continuous proportion; range 0 to 1, where 0= has sex without condoms 0% of the time and 1= has sex without condoms 100% of the time) reported by participant *i* at the six-month post intervention follow-up.

Y_{Pre} – The baseline measure of the outcome variable, inconsistency of condom use, reported by participant at baseline; variable re-centered at the grand mean for analysis.

T –A treatment indicator variable whose value equals 1 if the participant was randomized into the treatment group and zero otherwise.

X – A *p* vector of baseline participant-level covariates as well as blocking variables to account for the variation in outcomes associated with these groups. These covariates include:

a) Age – self reported age at baseline (continuous; range 14-19); variable re-centered at the grand mean for analysis.
b) Race – self-reported race of participant. A set of 4-1 = 3 dummy variables; for each a participant is coded as 1 if she identified as the specified race and 0 otherwise (included in the model are Black, Multiracial, and Other; White is the reference variable); variable re-centered at the grand mean for analysis.
c) Ethnicity – self-reported ethnicity of participant. A dummy variable (0= not Hispanic or Latino; 1=Hispanic or Latino); variable re-centered at the grand mean for analysis.
d) Parental education – A continuous measure of the mean level of parents' education reported by participants (scores range from 1 = less than high school to 5 = graduate degree); variable re-centered at the grand mean for analysis.
e) Family structure – A dummy indicator variable that measures whether or not a respondent lives with both parents (0= does not live with both parents; 1= lives with both parents); variable re-centered at the grand mean for analysis.
f) Site – a set of 4 dummy variables to capture the variable effects of the 5 clinic sites that offered the interventions during the evaluation period. For each variable, an individual participant was coded as 1 if she was randomly assigned at that particular site and 0 otherwise (Included in the model are Clinic A, Clinic B, Clinic C, and Clinic E; Clinic D is the reference variable). Dummy variables grand mean centered for analysis.

β_0 – The intercept term, which represents the mean self-reported inconsistency of condom use for control participants, six months after the end of treatment, with all other variables in the model held constant at zero.

β_1 – This is the parameter estimate of substantive interest. β_1 represents the adjusted mean difference in treatment and control participants' self-reported inconsistency of condom use six months after the end of treatment.

Data Cleaning Procedures

In order to and improve the validity and reliability of our estimates, prior to analysis, we followed several steps to prepare the dataset and improve the quality of our data. These procedures are outlined below.

Identify and flag unit and item non-response

The first step in the data screening process was to determine whether or not any individuals in the randomized sample had no data at all (i.e., did not complete a questionnaire) at baseline or at the six-month follow up observations. These cases were flagged as unit-nonresponse and treated as unit missing according to our missing data approach (see Missing Data Approach section below). Next, we identified instances in which no response was provided to a particular questionnaire item. These cases were flagged as item-nonresponse and treated as item-missing according to our missing data approach.

Identify and flag unreliable cases.

The second step was to identify cases (i.e., units or entire questionnaires) that were unreliable. By unreliable, we mean that we have sufficient reason to believe that the respondent's answers were not honest representations of their behaviors, knowledge, and beliefs. Cases were flagged as unreliable for three reasons: responses followed a clear, deliberate pattern; respondents finished the questionnaire in a time considered too fast to have read the questions and provided reliable responses (7 minutes or less for online questionnaires; 10 minutes or less for paper questionnaires); or the respondents indicated on their questionnaires that they were not honest as they responded. For our benchmark analyses, unreliable data were treated as unit missing (i.e., we considered the entire questionnaire as missing) and excluded from benchmark analyses. However, sensitivity analyses that included the unreliable data were conducted and results are presented in Appendix E.

Identify and flag invalid responses

The third step in the data screening process was to inspect the data for instances in which responses were invalid because they were outside of a pre-determined range of plausible or acceptable values. Each questionnaire has a codebook, which was prepared by PRG staff, that contains variable names and valid variable values or ranges of values. For each item, response values that were outside of the range specified in the codebook were flagged as invalid; these cases were treated as item missing.

Identify and flag outliers.

The fourth step was to identify and flag severe outliers. By outliers, we are referring to values that are extreme compared to other observations, but are not invalid. The only items for which we inspected outliers are those used in the construction of our outcome variables (see Tables III.1 and III.2) because they have no upper limit; therefore, responses could technically approach infinity (all other variables used in analysis are either categorical or have predicated upper and lower bounds). We defined values as severe outliers according their relation to the interquartile range (IQR). Severe outliers are those values outside of the *outer fences* of the population distribution.

Our benchmark analytic approach was to include data flagged as outliers in analysis, because we did not know for certain whether the values are true or invalid. However, we also ran sensitivity analyses that treated these data as item missing, and we report results in Appendix E.

Identify and flag inconsistencies in reporting of sexual behaviors.

The final step in the data review process was to inspect the data and identify inconsistencies in sexual behavior outcome data. With repeated measures of sexual behaviors, two primary types of inconsistencies occur – internal inconsistences and over-time inconsistencies. Internal inconsistencies refer to discrepancies in responses (to related questions) in the same survey administration. For instance, a respondent might say that she has not had sex in the past three-months, but then indicates that she used condoms three of the times she had sex in the past three-months. Over-time inconsistencies refer to instances in which lifetime reported behaviors decline or are completely recanted over time. For example, at baseline a respondent might say that she has had sex 10 times in her life, but on the subsequent administration of the survey she says either a) she has never had sex, or b) she has sex four times in her life.

Variables used in the construction of outcome variables were flagged as inconsistent data in the following instances:

Inconsistent internally

If, on one questionnaire (baseline or six-month follow-up), a respondent indicates that she has had sex in the past three months (i.e., she provides a response greater than "0" to one of the following questions: "*In total, how many times have you had any type of sex in the past 3 months?"*; "*In total, how many times have you had any type of sexual intercourse in the past 3 months?"*) but then indicates in the same survey administration that she has never had sex (i.e., she responds "*no*" to the question, "*Have you ever had any type of sex?")* all sexual behavior responses are flagged as inconsistent internally and recoded to missing.

If, on one questionnaire (baseline or six-month follow-up), a respondent indicates that she has not had sex in the past three-months (i.e., she responds "0" to the question, "*In total, how many times have you had any type of sex in the past 3 months?*), but then indicates in the same survey administration that she has used condoms while having sex in the past three months (i.e., she provides a response greater than "0" to the question, "*Now, think about the number of times that you had any type of sex in the past 3 months. How many of those times did you use condoms?"*), both responses are flagged as inconsistent internally and recoded to missing.

If, on one questionnaire (baseline or six-month follow-up), a respondent indicates that she has not had sexual intercourse in the past three months (i.e., she responds "0" to the question, "*In total, how many times have you had sexual intercourse in the past 3 months?"*) but then indicates in the same survey administration that she has had sexual intercourse without using birth control in the past three months (i.e., she provides a response greater than "0" to the question, "*In the past 3 months, how many times have you had sexual intercourse <u>without</u> using any of these methods of birth control?"),* both responses are flagged as inconsistent internally and recoded to missing

If, on one questionnaire (baseline or six-month follow-up), a respondent indicates that she has used condoms more times in the past three months than she has had sex (i.e., her response to the question, "*Now, think about the number of times that you had any type of sex in the past 3 months. How many of those times did you use condoms?"* is greater than the response given to the question, "*In total, how many times have you had any type of sex in the past 3 months?"*), both responses are flagged as inconsistent internally and recoded to missing.

If, on one questionnaire (baseline or six-month follow-up), a respondent indicates that she has not used birth control during sexual intercourse more times in the past three months than she has had sexual intercourse (i.e., her response to the question, "*In the past 3 months, how many times have you had sexual intercourse without using any of these methods of birth control (options listed)?*" is greater than the response given to the question, "*In total, how many times have you had any type of sexual intercourse in the past 3 months?*"), both responses are flagged as inconsistent internally and recoded to missing.

Inconsistent over time

If, at baseline, a respondent indicates that she has had sex during the past three months (i.e., she provides a response greater than "0" to the question, "*In total, how many times have you had any type of sex in the past 3 months?*), then, at six-month follow-up indicates that she has never had sex (i.e., she responds, "*I have never had any type of sex*" to the question, "*How old were you the first time you had any type of sex?"*), outcome measures at both baseline and follow-up are flagged as inconsistent over time and recoded to missing.

If, at baseline, a respondent indicates that she has had sexual intercourse during the past three months (i.e., she provides a response greater than "0" to the question, "*In total, how many times have you had sexual intercourse in the past 3 months?*), then, at six-month follow-up, indicates that she has never had sex (i.e., she responds "*no*" to the question, "*Have you ever had sexual intercourse?"* or *"Have you ever had any type of sex"*), outcome measures at both baseline and follow-up are flagged as inconsistent over time and recoded to missing.

Results of Data Cleaning

Below, in Table D.1, we present the results of our data cleaning procedures. For each variable used in analysis and in the construction of our outcome measures, we outline the number of observations that were treated as item-missing in our benchmark analyses due to nonresponse, to invalid response, or inconsistent responses. We also indicate the number of observations which were flagged as outliers (but not treated as item-missing in our benchmark analysis). In addition to what is reported below, 37 cases in our randomized sample were flagged as unit-nonresponse (2 did not complete a baseline questionnaire and 35 did not complete a six-month follow-up), and 11 were flagged as unreliable and treated as unit missing in our benchmark analysis.

Table D.1. Number of observations flagged as item non-response, invalid, inconsistent or outliers

Variable	Non-response	Invalid	Inconsistent	Outlier
Age	0	0	n.a	n.a.
Race	1	0	n.a	n.a
Ethnicity (Hispanic)	0	0	n.a	n.a
Parental education	1	18[a]	n.a	n.a
Family structure	2	0	n.a	n.a
Frequency of sex: baseline[b]	28	2	19	13
Frequency of condom use: baseline[c]	26	1	8	16
Frequency of sexual intercourse: baseline[d]	10	7	15	14
Frequency of sexual intercourse without birth control: baseline[e]	148	1	3	8
Frequency of sex: 6-month followup	36	2	19	16
Frequency of condom use: 6-month followup	45	1	13	16
Frequency of sexual intercourse: 6-month followup	1	4	7	14
Frequency of sexual intercourse without birth control: 6-month followup	132	2	1	13

Notes: n.a means not applicable [a]Though the response option, "I don't know" was provided for questions related to parents' educational attainment, these responses were treated as invalid as they could not be recoded to a meaningful value. [b]Frequency of sex refers to the number of times in the past three months a person reports having any type of sex (anal, oral, vaginal). This variable is used to assess our secondary research question related to the frequency of sex and it is used along with the variable "frequency of condom use" to construct the measure "inconsistency of condom use" for our primary research question. [c] Frequency of condom use refers to the number of times in the past three months a person reports using condoms while having any type of sex (anal, oral, vaginal). [d]Frequency of sexual intercourse refers to the number of times in the past three months a person reports having sexual intercourse (vaginal sex). This variable is used along with the variable "frequency of sexual intercourse without birth control" to construct the measure for our secondary research question related to inconsistency of contraceptive use. [e]Frequency of sexual intercourse without birth control refers to the number of times in the past three months a person reports having sexual intercourse (vaginal sex) without the use of any contraceptives (including condoms).

Missing Data Approach

Our missing data approach distinguishes between two types of missing data – *unit missing* and *item missing* – which are treated differently for the purposes of analysis. Unit missing refers to cases in which an entire questionnaire was not completed (e.g., a respondent failed to submit a follow-up questionnaire) or in which responses on a questionnaire were deemed unreliable (see Data Cleaning section above). Item missing refers to cases in which a questionnaire was completed and considered reliable; however, there was no response provided for one or more items on the questionnaire or (for one or more item) responses provided were invalid or inconsistent and were recoded to missing. The benchmark approach to missing data that we selected takes both types of missing data into account and aims to mitigate the introduction of bias into our impact estimates, provide good estimates of uncertainty, and maximize the use of available data by imputing or adjusting data.

Our six-step decision process is outlined below.

1. Using data cleaning procedures outlined in the *Data cleaning* section, identify inconsistent, unreliable, and invalid data in any analytic (i.e., outcome, pretest, or covariate) variables and recode inconsistent and invalid data as item missing and flag unreliable data as unit missing for analysis.
2. Examine prevalence of data flagged as unit and item missing for both treatment and control samples.
3. Determine if logical imputations are possible for any analytic variables that may have missing values (due to nonresponse) and logically impute where this is the case.
4. Drop from the analytic sample any participants for whom at least one questionnaire (baseline or six-month follow-up) is considered unit missing. We reasoned that case-wise deletion is the most prudent approach, as no reliable data exist at the individual-level from which to estimate values for the missing data.
5. For the remaining missing analytic data we then imputed or adjusted the missing values differently depending on whether the variables are: (a) pretest (and other covariate) data, or (b) posttest or outcome data.
    a. For missing pretest or covariate data, our benchmark approach is to use dummy variable adjustment procedures. While Puma et al. (2009) concede that this approach is questioned in the literature, they recommend it as a preferred approach regardless of whether data are missing at random, missing completely at random or missing not at random. They argue and find in their simulations that it is an appropriate strategy to maximize the analytic sample without biasing results as long as the assignment to treatment is uncorrelated with the covariate missing data (which it should be, given that random assignment ensures that treatment is in expectation exogenous and unrelated to all observed covariates).
    b. For missing posttest data, our benchmark approach is to use Multiple Stochastic Regression Imputation. Puma et al. (2009) recommend this as one approach that minimizes bias in their simulations. Briefly, this is a regression-based approach to imputation that imputes missing values with predicted values derived from the combination of multiple (in our case 10) iterations of the dataset (i.e., 10 separately constructed datasets with distinct predicted

values). With this approach, variance is to be the same across imputed and observed values. Note that we did not to impute separately by treatment condition; our benchmark approach was to include the treatment indicator as a variable in the imputation model. However, we did run separate analyses in which we imputed separately by condition and the results were substantively the same (the regression adjusted mean differences between groups were in the same direction and none were statistically significant at the $\alpha = .05$ level).

In Table D.2, below, we present the number of observations in our analytic sample that were treated as item-missing and imputed in our benchmark analysis, by treatment group status.

Table D.2. Number of item-missing observations for each analytic variable, by treatment status and full sample

Variable	Safer Sex Intervention	Female Sexual Health	Full analytic sample
Age	0	0	0
Race	1	0	1
Ethnicity: Hispanic	0	0	0
Parental education	10	9	19
Family structure	1	1	2
Frequency of sexual activity: baseline	25	24	49
Inconsistency of condom use: baseline	25	28	53
Inconsistency of contraceptive use: baseline	80	83	163
Frequency of sexual activity: 6-month follow-up	23	34	57
Inconsistency of condom use: 6-month follow-up	24	35	59
Inconsistency of contraceptive use: 6-month follow-up	71	69	140
Sample size	133	135	268

6. The final step of the decision making process was to conduct sensitivity analyses by estimating results with missing data excluded from the analysis (i.e., use case-wise deletion for all cases with missing data in analytic variables). In Appendix E, we report our benchmark results next to the sensitivity analysis results to verify findings. In addition, because the discussion of our analytic sample and baseline equivalence results presented previously in this report reflect data that are partially imputed, below in Table D.3, we present results of baseline equivalence tests that use unimputed data. Results are similar to those previously reported, in no case are significant differences between the treatment and control group apparent (p-values > .05).

Table D.3. Summary statistics of key baseline measures, item-missing data are not imputed

Baseline measure	Sample size	Intervention mean or proportion(sd)	Comparison mean or proportion (sd)	Mean difference	*p*-value of difference
Age (in years)	268	17.42 (1.35)	17.65 (1.29)	-0.23	0.14
Race: White	267	0.11 (0.32)	0.08 (0.27)	0.03	0.41
Race: Black	267	0.79 (0.41)	0.83 (0.38)	-0.04	0.42
Race: Multiracial[a]	267	0.07 (0.27)	0.07 (0.25)	0.01	0.82
Race: Other[b]	267	0.02 (0.15)	0.02 (0.15)	0.00	0.97
Ethnicity: Hispanic	268	0.02 (0.15)	0.07 (0.25)	-0.04	0.08
Parental education[c]	249	2.31 (1.07)	2.13 (0.92)	0.18	0.13
Family structure (lives with both parents)		0.10 (0.31)	0.08 (0.28)	-0.02	0.55
Frequency of sexual activity[d]	219	13.20 (20.02)	13.59 (20.67)	-0.4	0.89
Inconsistency of condom use[e]	215	0.43 (0.4)	0.47 (0.44)	-0.04	0.52
Inconsistency of contraceptive use[f]	105	0.47 (0.41)	0.47 (0.41)	0	0.99

Appendix E: Sensitivity analyses

Study 1: Baseline Covariates

We test our benchmark approach of including covariates (including the baseline measure of the outcome variable) in the analytic model by estimating an otherwise identical empirical model without the covariates included and comparing the sensitivity model estimates with the benchmark model estimates. Coefficients and p-values for the treatment indicators for the two contrasts are presented in tables E.1 and E.2 below under Study 1. The estimates produced by both models are substantively identical; results indicate no programmatic effect on inconsistency of condom use, inconsistency of contraceptive use, or frequency of sex six months after the treatment intervention ends. The *p*-values are considerably greater than 0.05 for all three outcomes. Consequently, we infer that substantive findings are identical regardless of whether or not we control for covariates in the analytic model.

Study 2: Missing Data

As detailed in the *Impact Analysis Plan* (replicated in the attached Appendix D: Methods), we specify a benchmark approach that relies on imputation and adjustment of data to reduce attrition in our analytic sample. We test this approach by comparing benchmark results with those produced by the same empirical model but with a reduced analytic sample that does not include cases that rely on imputed or adjusted data (n = 167, n = 59, and n = 169 for inconsistency of condom use, inconsistency of contraceptive use, and frequency of sex, respectively). Coefficients and *p*-values for the treatment indicator are presented in the tables below under Study 2. Again, as can be seen in Tables E.1 and E.2 the results produced with both analytic samples do not change inferential findings. The estimated treatment effects for both the benchmark and alternative reduced sample are not significant. Consequently, we infer that findings are not sensitive to the decision to impute or otherwise adjust missing data as detailed in the *Impact Analysis Plan*.

Study 3: Unreliable Data

In our benchmark analytic approach, we treat cases with what is deemed to be unreliable data as unit missing (see data cleaning section in Appendix D) and exclude them from the analytic sample (n = 268). We test whether or not this analytic decision has an effect on substantive findings by comparing benchmark results with those produced by the same procedures, but with an analytic sample that includes the cases with unreliable data (n = 279). Estimated treatment effects (coefficients and *p*-values) for this analytic sample are presented in Tables E.1 and E.2 tables under Study 3. As can be seen, the results are inferentially similar. Estimated treatment effects are statistically insignificant for all models and outcomes.

Study 4: Outliers

Our benchmark approach is to include all cases with observations that are identified as outliers (see *data cleaning* section of Appendix D). We test whether or not this analytic decision has an effect on inferential findings by comparing benchmark results with those produced by the same procedures converting all outliers to missing and imputing as we would other missing or invalid data. Coefficients and *p*-values for the treatment indicator are presented in the tables below under Study 4. As can be seen, the inclusion of outliers does not change inferential findings. For all three outcomes, the coefficients for the treatment indicators are not significant in either the benchmark or sensitivity data.

Study 5: Conflicting Eligibility Data

To be considered eligible for enrollment, during the screening process, potential participants must have reported that they were between 14 and 19 years old, that they had engaged in sex with a male in the three months prior, and that they had not participated in a specified list of other OAH funded TPP programs operating in the New Orleans area. If participants reported that they met these criteria and they provided consent/assent (prior to their initial session), they were randomized and enrolled in the study.

However, in a number of cases, participants' responses on the subsequent baseline questionnaire conflicted with information provided at screening. Specifically, in 11 instances, participants' self-reported birthdays suggest they were at least 20 years old at baseline (on average, screening and baseline occurred within six to seven days of each other). In 8 instances participants indicated that they had not had sex with a male in the prior three months, and, in 48 instances participants indicated that they had participated in an OAH TPP funded program. Our benchmark approach was to include these 67 individuals in our analyses since they were randomized, and we had no way of confirming which reports are accurate. However, we also test whether or not the inclusion of these individuals' responses changes the substantive interpretation of our results by conducting a sensitivity study in which we exclude these cases (n = 201). Coefficients and *p*-values for the treatment indicator are presented in the tables below under Study 5. As can be seen, the exclusion of individuals with conflicting eligibility data of does not change findings. For all three outcomes, the coefficients for the treatment indicators are insignificant in all the analyses conducted with the benchmark and sensitivity samples.

Table E.1. Sensitivity of impact analyses using data collected 6 months post treatment to address the primary research questions

Intervention compared to control	Benchmark *b*	Benchmark *p*	Study 1 *b*	Study 1 *p*	Study 2 *b*	Study 2 *p*	Study 3 *b*	Study 3 *p*	Study 4 *b*	Study 4 *p*	Study 5 *b*	Study 5 *p*
Inconsistency of condom use	0.04	0.642	0.04	0.676	0.05	0.429	0.05	0.620	0.02	0.774	0.08	0.453

Source: Six-month follow-up survey.

Notes: *b* refers to the regression adjusted mean difference in the outcome between Safer Sex Intervention and Female Sexual Health. *p* refers to the p-value of the difference; results are considered significant if $p < .05$. See Table III.3 for a more detailed description the outcome measure and section III for a description of the impact estimation methods.

Table E.2. Sensitivity of impact analyses using data collected 6 months post treatment to address the secondary research questions

Intervention compared to control	Benchmark *b*	Benchmark *p*	Study 1 *b*	Study 1 *p*	Study 2 *b*	Study 2 *p*	Study 3 *b*	Study 3 *p*	Study 4 *b*	Study 4 *p*	Study 5 *b*	Study 5 *p*
Inconsistency of contraceptive use	0.15	0.254	0.14	0.302	0.04	0.682	0.18	0.192	0.17	0.164	0.21	0.370
Frequency of sex	3.72	0.417	4.15	0.391	2.19	0.492	4.11	0.390	-0.64	0.763	1.15	0.847

Source: Six-month follow-up survey.

Notes: *b* refers to the regression adjusted mean difference in the outcome between Safer Sex Intervention and Female Sexual Health. *p* refers to the p-value of the difference; results are considered significant if $p < .05$. See Table III.3 for a more detailed description the outcome measure and section III for a description of the impact estimation methods.

Appendix F: Implementation evaluation methods

Table F.1. Methods used to address implementation research questions

Implementation element	Methods used to address each implementation element
Adherence to Safer Sex Intervention program model: How many sessions were offered?	Total number of initial, 1-month booster, 3-month booster, and 6-month booster sessions offered is a sum of the session type offered captured by the Enrollment Log; reported overall and by study site.
Adherence to Safer Sex Intervention program model: What and how much was received?	Percentage of study participants attending each session type is calculated as the total number of each session type attended divided by the total number of each session type offered as captured by the Enrollment Log; reported overall and by study site.
	Average session duration is calculated as the sum of the total number of minutes of each session attended (for initial, 1-, 3-, and 6-month boosters separately) divided by the total number of sessions attended. Both the numerator and denominator are captured by the Enrollment Log.
	Average session time for participants who completed one, two, three, and four sessions is calculated as the sum of the total number of minutes for participants completing each session combination (initial only, initial plus one booster, initial plus two boosters, initial plus three boosters) divided by the total number of participants completing the session combination.
	Percentage of treatment sample that did not attend any sessions is calculated as the total number of treatment participants who failed to attend any session divided by the total number of treatment participants.
	Percentage of treatment sample that attended all sessions is calculated as the total number of treatment participants who attend all four sessions divided by the total number of treatment participants.
	Percentage of treatment sample that attended all sessions within the intended time frame is calculated as the total number of treatment participants who attend all four sessions within 60-140 minutes divided by the total number of treatment participants.
	Average number of sessions attended per participant is calculated as the sum of the total number of sessions attended for each treatment participant divided by the total number of treatment participants. (Note: a participant may attend a maximum of four sessions: initial, 1-month booster, 3-month booster, and 6-month booster.)
	Percentage of initial sessions completed within and outside of the intended session duration range is calculated as the number of sessions completed that lasted 1) between 30-50 minutes, 2) less than 30 minutes, and 3) more than 50 minutes, divided by the total number of initial sessions completed.
	Percentage of booster sessions completed within and outside of the intended session duration range is calculated as the number of 1-, 3-, and 6-month booster sessions completed that lasted 1) between 10-30 minutes, 2) less than 10 minutes, and 3) more than 30 minutes, divided by the total number of booster sessions completed.
Adherence to Safer Sex Intervention program model: What amount of content was delivered to youth?	Average number of intervention activities completed for each session type is calculated as the sum of the total number of activities completed (for initial pre-contemplation, initial contemplation, 1-, 3-, and 6-month boosters separately) divided by the total number of sessions for which we have health educator self-reports/fidelity monitor observations, reported for each session type. (Note: There are 11 initial session pre-contemplation stage activities, 12 initial session contemplation stage activities, and 6 booster session activities at the 1-, 3-, and 6-month booster sessions. An activity is considered complete if all components within that activity have been marked as completed by the health educator/fidelity monitor observer on the Safer Sex Intervention Fidelity Toolkit Forms. If any component in an activity is marked as not applicable, we interpret that activity as not completed.)

Implementation element	Methods used to address each implementation element
	Percentage of initial sessions in which 100% of intervention activities are completed is calculated as the sum of the total number of each session type (initial pre-contemplation, initial contemplation, 1-month, 3-month, and 6-month booster) in which the intended number of activities were completed, divided by the total number of sessions for which we have health educator self-reports/fidelity monitor observations.
Adherence to Safer Sex Intervention program model: Who delivered material to youth?	Total number of staff who delivered the program (overall and by study site) is a sum of all health educators who facilitated at least one Safer Sex Intervention session, as captured by the Enrollment Log.
	List of each health educator's employment status (intern, part-time, or full-time employee) and credentials (degree/certifications) (reported overall and by site). (Note: each health educator was assigned a code, which was used when reporting statistics.)
	Percentage of staff trained in Safer Sex Intervention, Female Sexual Health control condition, Fidelity Monitoring Procedures, and SMARTS Research Protocols is calculated as the number of staff members who completed all four trainings divided by the total number of staff who delivered the program.
Quality: Quality of staff-participant interactions (Each statistic will be reported by site and overall, for treatment groups only.)	Percentage of observed sessions where the fidelity monitor scored the **delivery of session information** to participants as "good" (=4) or "very good" (=5) is calculated as the total number of sessions for which the average score for questions 1-3 from the Program Observation Form for TPP Grantees = 4 or 5, divided by the total number of observed sessions. (Delivery of session information is a scale variable that is constructed as the average score of item responses to questions 1-3 in the Program Observation Form for TPP Grantees. Response options for questions 1-3 range from 1-5, with 1 being the worst rating and 5 being the best rating. For each rated session, the scale score could range from 1=very poor delivery of information to 5=very good delivery of information).
	Percentage of observed sessions where the fidelity monitor scored the **extent of participants' understanding** of session material as "moderate" (=4) or "good" (=5) is calculated as the total number of sessions for which the score for question 4 in the Program Observation Form for TPP Grantees = 4 or 5, divided by the total number of observed sessions. (Extent of participants' understanding is operationalized as the response to question 4 in the Program Observation Form for TPP Grantees; response items range from 1=little understanding to 5=good understanding.)
	Percentage of observed sessions where the fidelity monitor scored the **level of participation** in session discussions and activities as "moderate" (=4) or "active" (=5) is calculated as the total number of sessions for which the score for question 5 in the Program Observation Form for TPP Grantees = 4 or 5, divided by the total number of observed sessions. (Level of group participation is operationalized as the response to question 5 from Program Observation Form for TPP Grantees; response items range from 1=little participation to 5=active participation.)
	Percentage of observed sessions where the fidelity monitor scored the **overall quality of the program session** as "very good" (=4) or "excellent" (=5) is calculated as the total number of sessions for which the score for question 7 in the Program Observation Form for TPP Grantees = 4 or 5, divided by the total number of observed sessions. (Overall quality of program session is operationalized as the response to question 7 from Program Observation Form for TPP Grantees; response items range from 1=poor to 5=excellent.)
Counterfactual comparison condition experiences: What was offered?	Total number of sessions offered is a sum of the sessions offered captured by the Enrollment Log.
Counterfactual comparison condition experiences: What was received?	Percentage of sessions attended is calculated as the total number of sessions attended divided by the total number of sessions offered as captured by the Enrollment Log.

Implementation element	Methods used to address each implementation element
	Average counterfactual session duration is calculated as the sum of the total number of minutes of each session attended divided by the total number of sessions attended. Both the numerator and denominator are captured by the Enrollment Log.
	Average number of sessions attended is calculated as the sum of the total number of sessions attended divided by the total number of sessions offered. Both numerator and denominator are captured by the Enrollment Log.
Counterfactual comparison condition experiences: What amount of content was delivered to youth?	Average number of counterfactual intervention topics on which information was provided to participants is calculated as the sum of the total number of topics presented at each counterfactual session divided by the total number sessions for which we have health educator self-reports/fidelity monitor observations. (Note: The counterfactual intervention, Female Sexual Health, has 16 topics that should be presented to participants. A topic is considered complete if it has been marked as provided by the health educator/fidelity monitor on the Fidelity Monitor Self-Report/Observer Form.)
	Percentage of counterfactual sessions in which 100% of intervention topics are provided is calculated as the total number of sessions in which 16 topics are provided divided by the total number of sessions for which we have health educator self-reports/fidelity monitor observations.
	Percentage of counterfactual sessions in which the health educator engaged in any (one or more) of the five assessed components of Safer Sex Intervention (the treatment condition) is calculated as the total number of sessions in which the health educator engaged in 1 or more components divided by the total number of sessions for which we have health educator self-reports/fidelity monitor observations. (Note: There are five Safer Sex Intervention components assessed on the fidelity monitoring forms.)
Counterfactual comparison condition experiences: Who delivered material to youth?	Analysis is the same as is listed in the adherence section.
Context: Other TPP programming available or offered to study participants (both intervention and counterfactual)	A list of all other TPP programming being implemented in Orleans Parish during the program period (and thus potentially available to both intervention and comparison groups).
	Percentage of participants self-reporting past-year exposure to reproductive health education is calculated as the total number of participants who report past-year exposure to reproductive health education divided by the total number of participants who complete the questionnaire. (This statistic is reported by treatment and comparison group for baseline and six-month survey administrations.)
	Percentage of participants self-reporting past-year experiences with other TPP programs is calculated as the total number of participants who report past-year experiences with other TPP programs divided by the total number of participants who complete the questionnaire. (This statistic is reported by treatment and comparison group for baseline and six-month survey administrations)

Implementation element	Methods used to address each implementation element
Context: External events affecting implementation	A list of external events that did or may have affected program implementation.
Context: Substantial unplanned adaptation(s)	A list of any substantial unplanned adaptations to the program, for which adaptation requests were made to OAH.

TPP = Teen Pregnancy Prevention.

Appendix G: Implementation Evaluation Results

ADHERENCE TO PROGRAM MODEL

How Many Sessions Were Offered

Table G.1. Total number of Safer Sex Intervention sessions offered, by study site and overall

Study site	Initial session	1-month booster	3-month booster	6-month booster
Clinic A	24	24	24	24
Clinic B	14	14	14	14
Clinic C	44	44	44	44
Clinic D	72	72	72	72
Clinic E	5	5	5	5
Overall (sum)	159	159	159	159

What and How Much Was Received

Table G.2. Percentage of study participants attending each Safer Sex Intervention session, by study site and overall

Study site	Number of participants	Initial session	1-month booster	3-month booster	6-month booster
Clinic A	24	95.8	20.8	25.0	25.0
Clinic B	14	100.0	50.0	42.9	35.7
Clinic C	44	100.0	70.5	79.5	75.0
Clinic D	72	98.6	86.1	81.9	63.9
Clinic E	5	100.0	0.0	20.0	0.0
Overall (sum)	159	98.7	66.0	67.3	56.6

Table G3. Average number of Safer Sex Intervention sessions attended, by study site and overall

Study site	Number of participants	Average number of sessions
Clinic A	24	1.7
Clinic B	14	2.3
Clinic C	44	3.3
Clinic D	72	3.3
Clinic E	5	1.2
Overall (sum)	159	2.9

Notes: A participant may attend a maximum of four sessions: initial, one-month booster, three-month booster, and six-month booster.

Table G.4. Percentage of study participants who attended no intervention sessions, all intervention sessions, and all sessions within intended time frame

	Number of participants	Percent of participants
Participants who attended no sessions	159	1.3
Participants who attended all (4) sessions	159	36.5
Participants who attended all (4) sessions within intended time frame [a]	159	27.7

Notes: A participant may attend a maximum of four sessions: initial, one-month booster, three-month booster, and six-month booster. [a] Intended length of initial session is 30-50 minutes, and intended length of each booster session is 10-30 minutes; therefore, the range of time for the intended duration for a participant attending all four sessions would be 60-140 minutes. Only participants who attended all four sessions *and* for whom we have session duration data for all four sessions are included in this analysis.

Table G.5. Average Safer Sex Intervention session duration (in minutes), by session type

Session type	Average time (in minutes)	Intended duration or range (in minutes)
Initial session	49.8	30-50
1-month booster	8.6	10-30
3-month booster	9.5	10-30
6-month booster	8.4	10-30

Table G.6. Percent of Safer Sex Intervention initial and booster sessions completed within and outside of intended session duration range

	Sample size	Percent of sessions
Initial sessions completed within 30-50 minutes	153	68.6
Initial sessions less than 30 minutes	153	5.9
Initial sessions more than 50 minutes	153	25.5
Booster sessions (1, 3, or 6) completed within 10-30 minutes	294	47.6
Booster sessions (1, 3, or 6) less than 10 minutes	294	51.4
Booster sessions (1, 3, or 6) more than 30 minutes	294	1.0

Notes: A participant may attend a maximum of four sessions: initial, one-month booster, three-month booster, and six-month booster. Intended length of initial session is 30-50 minutes, and intended length of each booster session is 10-30 minutes.

Amount of Content Delivered to Youth

Table G.7. Average number of intervention activities completed for each Safer Sex Intervention session

	Average number of activities completed	Intended number of activities
Initial pre-contemplation	9.8	11
Initial contemplation	9.8	12
1-month booster	4.9	6
3-month booster	4.9	6
6-month booster	4.9	6

Notes: Data on number of intervention activities completed in each session type comes from health educator self-reports and fidelity monitor observer reports, which were only completed for a subsample of intervention sessions; we use observer reports if they exist, otherwise we use health educator self-report data. An intervention activity is only considered complete if all components within that activity were marked as completed by the health educator/fidelity monitor observer on the Safer Sex Intervention Fidelity Toolkit Forms. If any component of an activity was marked as "not applicable", we interpret that activity as not completed.

For activity 1 of the pre-contemplation and contemplation sessions, there are two components listed on the Safer Sex Intervention *Fidelity Toolkit Self-Report and Observer Forms*: 1) developed rapport between participant and educator through introduction and discussion of confidentiality and goals and 2) showed the first segment of "Breaking Out" of the "Private Lives: STI and HIV Education" video. For unknown reasons, fidelity data were only entered into the data system for the *second component* of activity 1 (watching the video); therefore, we will consider activity 1 completed in the above and subsequent tables if the datasets indicate that this component was conducted.

Limitations: 1) health educator self-reports may not be a reliable measure of the content that was actually delivered to participants; additionally, we do not have complete self-report data for all Safer Sex Intervention sessions delivered; we have self-report data for: 89.2% (140/157) of initial sessions, 82.9% (87/105) of one-month booster sessions, 66.4% (71/107) of three-month booster sessions, and 60.0% (54/90) of six-month booster sessions; 2) fidelity monitor observer data are very incomplete and may thus fail to offer a representative picture of the content actually delivered to youth; we have limited observation data for all Safer Sex Intervention session types - we have observation data for: 15.3% (24/157) of initial sessions, 5.7% (6/105) of one-month booster sessions, 0.9% (1/107) of three-month booster sessions, and 0.0% (0/90) of six-month booster sessions.

Table G.8. Percentage of Safer Sex Intervention sessions in which 100% of activities completed

	100% of activities completed	Intended number of activities
Initial pre-contemplation	64.3	11
Initial contemplation	14.4	12
1-month booster	13.6	6
3-month booster	7.0	6
6-month booster	5.6	6

Notes: Data on number of intervention activities completed in each session type comes from health educator self-reports and fidelity monitor observer reports, which were only completed for a subsample of intervention sessions; we use observer reports if they exist; otherwise, we use health educator self-report data. For pre-contemplation initial sessions, we consider 75% to be 8 or more activities and 100% to be 11 activities completed; for contemplation initial sessions, we consider 75% to be 9 or more activities and 100% to be 12 activities. For booster sessions, we consider 100% completion if all 6 activities are completed. An activity is only considered complete if all components within that activity have been marked as completed by the health educator/fidelity monitor observer on the Safer Sex Intervention Fidelity Toolkit Forms. If any component of an activity is marked as "not applicable", we will interpret that activity as not completed.

Limitations: see Table G.7

Who Delivered Material to Youth

Table G.9. Health educator credentials, employment status, and study sites

Health educator ID	Credentials	Employment status	Study sites health educator facilitated sessions
1	MSN, BA, RN	Full-time	Clinic A
2	BA, grad student	Full-time	Clinic A, Clinic E
3	MPH	Full-time	Clinic A, Clinic C
4	BS	Part-time	Clinic A
5	MPH	Full-time	Clinic B, Clinic C
6	BS	Full-time	Clinic A, Clinic B, Clinic C
7	BS, grad student	LPHI intern	Clinic C, Clinic D
8	MSW, grad student	Part-time	Clinic A, Clinic E
9	BS	Full-time	Clinic C, Clinic D
10	BA	Full-time	Clinic D
11	BA, grad student	LPHI intern	Clinic D

Notes: Staff background information is self-reported. Each health educator has been assigned an ID for reporting statistics. Seven health educators facilitated sessions at multiple study sites. The same individuals facilitated both the Safer Sex Intervention and the counterfactual condition, Female Sexual Health.

Not listed is a staff member at Clinic A who was not authorized or trained to conduct the interventions, but who facilitated the initial intervention with 13 girls enrolled at that site.

QUALITY OF STAFF-PARTICIPANT INTERACTIONS

Table G.10. Percentage of observed initial Safer Sex Intervention sessions in which staff-participant interactions were rated good/moderate or better by fidelity monitor observers, by study site and overall

Study site	Number of observations	Delivery of information (scored good/ very good)	Extent of participants' understanding (scored moderate/good)	Extent of participation (scored moderate/ active)	Overall quality of program session (scored good/excellent)
Clinic A	3	0.0	33.3	66.7	33.3
Clinic B	3	66.7	66.7	100.0	33.3
Clinic C	9	22.2	55.6	88.9	44.4
Clinic D	8	37.5	75.0	87.5	100.0
Overall	23	30.4	60.9	87.0	60.9

Notes: Data that are used to assess quality of staff-participant interactions are from the *Program Observation Form for TPP Grantees*. This form was only used to collect data for observed *initial* Safer Sex Intervention sessions; it is unknown why data were not collected for observed booster sessions. Data are not representative of all interactions; they are based on a limited convenience sample of observed sessions. We have *Program Observation Form* observation data for 14.6% (23/157) of Safer Sex Intervention initial sessions. Clinic E is not included in this table because only five study participants total were enrolled at this site, and no observer report data were collected for the sessions conducted there.

COUNTERFACTUAL

Sessions Offered and Received

Table G.11. Number of Female Sexual Health sessions offered, percentage of participants who attended session, average session duration (in minutes), and average number of sessions attended, by study site and overall

Study site	Number of participants	Number of sessions offered	Percent attending sessions offered	Average session duration (in minutes)
Clinic A	23	23	100.0	59.0
Clinic B	15	15	100.0	40.4
Clinic C	49	49	100.0	36.1
Clinic D	73	73	100.0	37.9
Clinic E	0	0	n.a.	n.a.
Overall	160	160	100.0	40.7

Notes: Female Sexual Health consists of one 30 to 50 minute session. However, six participants in the control group were offered and received a Safer Sex Intervention one-month booster session (4 at Clinic A; 1 at Clinic C; and 1 at Clinic D). The overall average session duration for the one-month boosters for these six participants was 19.8 minutes. n.a. means not applicable.

Amount of Content Delivered to Youth

Table G.12. Amount of Female Sexual Health session content delivered to participants

	Number of observations	Statistic
Average number of session topics completed	151	15.6
Percentage of sessions in which 100% of topics completed	151	84.1
Percentage of sessions in which health educator engaged in at least one Safer Sex Intervention component	151	9.3

Notes: The counterfactual session, Female Sexual Health, has 16 sexual health topics that should be presented to the participant. Data on number of session topics completed comes from health educator self-reports and fidelity monitor observer reports, which were only completed for a subsample of intervention sessions; we use observer reports if they exist, otherwise we use health educator self-report data. A topic is considered complete if it has been marked as provided by the health educator/fidelity monitor observer on the Female Sexual Health Fidelity Monitor Self-Report/Observer Form. Facilitator self-report and observer fidelity monitoring forms assessed if the facilitator engaged in any of the following five Safer Sex Intervention components: (1) motivational interviewing, (2) assessment of participant using Wheel of Change to identify how to customize session, (3) tailoring of session based on participant's feedback and stated priorities, (4) discussing consequences of sexual risk behavior and strategies to address them, and (5) teaching condom use and negotiation skills through demonstration and role-play. We consider a health educator as having engaged in Safer Sex Intervention components if they engaged in 1 or more of the five assessed components.

Limitations: 1) Health educator self-reports may not be a reliable measure of the content that was actually delivered to participants; additionally, we do not have complete self-report data for Female Sexual Health sessions delivered; we have self-report data for 90.0% (144/160) of sessions; 2) fidelity monitor observer data are very incomplete and may thus fail to offer a representative picture of the content actually delivered to youth; we have limited observation data for Female Sexual Health sessions, for just 17.5% (28/160) of sessions.

CONTEXT

Other TPP Programming Available or Offered to Study Participants (Both T and C)

Table G.13. List of other known teen pregnancy prevention programming being implemented in Orleans Parish during the program period

Program name	Program lead agency	Funder/ grantee type	City/State
Making Proud Choices (MPC!) (or: Believe in Youth! or BY!-NOLA!)	Institute of Women and Ethnic Studies	OAH - TPP-Tier 1	New Orleans, LA
Teen Outreach Program (TOPs Clubs)	Louisiana DHH Office of Public Health	OAH - TPP-Tier 1	New Orleans, LA
e-SiHLE	Tulane University	OAH - TPP-Tier 2	New Orleans, LA
Becoming a Responsible Teen (BART)	Louisiana Public Health Institute	OAH - TPP-Tier 1	New Orleans, LA
Project AIM (Adult Identity Mentoring) - adaptation	Louisiana Office of Public Health	OAH - PREP	New Orleans, LA
Focus on Your Future	University of Kentucky College of Public Health	NIMH – R01 Study	New Orleans, LA
Be Proud! Be Responsible! (BPBR)	*Central Louisiana Area Health Education Center Foundation*	*OAH - TPP-Tier 1*	*Alexandria, LA*
SiHLE	*Louisiana Office of Public Health*	*OAH – PREP*	*Regions outside of Orleans Parish*

Notes: BPBR and SiHLE are included at the bottom in italics in this table because they were implemented during the program period, but outside of Orleans Parish.

Table G.14. Percent of participants in analytic sample reporting past-year exposure to reproductive health education at each data collection point, overall and by treatment and control group

	Overall number	Overall percent	Safer Sex Intervention number	Safer Sex Intervention percent	Female Sexual Health number	Female Sexual Health percent
Baseline	256	51.6	129	53.5	127	49.6
6-month follow-up	260	56.2	129	58.9	131	53.4

Notes: Data presented in the table were obtained from one question on the *Safer Sex Program Questionnaire.* The question asks, "In the past year, please tell us if you have had any formal education classes in school or some other place, such as a community center, church, or health clinic, on any of the following: (Please choose ALL that apply)." The response options are: the female menstrual cycle (period); how pregnancy occurs; sexually transmitted infections (STIs); how to say "NO" to sex; methods of birth control – that is, how to stop a pregnancy from happening; how to prevent HIV/AIDS using safe sex practices; and, I have never had any formal educational classes on any of the above topics. Number refers to the number of participants in the analytic sample who responded to the question at each data collection point; percent refers to the percent of those responding who indicated they had been exposed to reproductive health education in the past year.

Table G.15. Percentage of participants in analytic sample self-reporting past-year experiences with one or more other teen pregnancy prevention programs at each data collection point, overall and by treatment and control group

	Overall number	Overall percent	Safer Sex Intervention number	Safer Sex Intervention percent	Female Sexual Health number	Female Sexual Health percent
Baseline	249	13.3	121	13.2	128	13.3
6-month follow-up	264	17.8	131	19.8	133	15.8

Notes: Data presented in the table were obtained from one question on the *Safer Sex Program Questionnaire.* The question asks "In the past year, have you been a participant in any of the following youth programs? (Please choose ALL that apply)." The response options are: Becoming a Responsible Teen (BART); Healthy Living; 4 Real Health; Be Proud! Be Responsible!; MPC! – NOLA (Making Proud Choice – New Orleans, LA); Teen Outreach Program (also known as TOPs Clubs); Safer Sex Intervention; SMARTS (Staying Mature and Responsible Toward Sex); Sisters Informing, Healing, Living, and Empowering (SiHLE); Project AIM (Adult Identity Mentoring); Focus on Your Future!; Other(s) – (write in option); I have never been a participant in any of the youth programs listed above. Number refers to the number of participants in the analytic sample who responded to the question at each data collection point; percent refers to the percent of those responding who indicated they had been exposed to reproductive health education in the past year.

Though SMARTS program names (Safer Sex Intervention, SMARTS) were included as response options for this question, participants who selected any of these three options were not counted as having an experience with any 'other TPP' program in the above table. However, it should be noted that although participants were diligently screened by study staff for prior participation in SMARTS before being enrolled in the study, at baseline, 25 participants (11 assigned to Safer Sex Intervention and 14 assigned to Female Sexual Health) self-reported on the questionnaire that they had participated in Safer Sex Intervention or SMARTS. While this is concerning, we recognize that self-reports are often unreliable and, though the question asked about their participation in these programs in the past year, it's possible that at least a portion of these youth answered affirmatively because they were currently enrolled in SMARTS (though they had not yet received their first program session at the time the baseline questionnaire was administered).