BOSTON CHILDREN'S HOSPITAL

EVALUATION OF THE MOMENTARY AFFECT REGULATION SAFER SEX INTERVENTION

OFFICE OF POPULATION AFFAIRS TEEN PREGNANCY PREVENTION PROGRAM

STATISTICAL ANALYSIS PLAN

NCT04798248

OCTOBER 2023

Submitted by:
The Policy & Research Group
www.policyandresearch.com

8434 Oak St.
New Orleans, LA 70118

107 Spring St.
Seattle, WA 98104

INTRODUCTION

OVERVIEW

In January 2023, The Policy & Research Group (PRG) submitted its Analysis Plan for the evaluation of the Momentary Affect Regulation Safer Sex Intervention (MARSSI) to Mathematica. In May 2023, Mathematica provided feedback to PRG on the plan.

In an effort to provide detail to other reviewers of this Analysis Plan (who may have similar questions), we include brief responses to key questions raised by Mathematica to offer further clarity on decisions made and how we intend to proceed with our analysis.

Reporting plan. In light of the number of research questions specified below, we want to be clear on how we intend to report findings from this evaluation. Our first priority will be to report results relevant to the three primary research questions using our benchmark analytic approach. This may include additional analyses, as needed, to clarify any analytical complications that invariably arise in the analysis of the impacts of complex social programs. We have noted some potential complexities in this Analysis Plan. Although it is not a requirement for an RCT with low attrition, we will include baseline equivalence statistics for each of the analytic samples used to produce findings for the three primary research questions. We may opt to include findings from secondary and exploratory research questions if we find that they elucidate or supplement the primary research findings.

We have asked an array of secondary and exploratory questions that aim to investigate hypothetical relationships that undergird MARSSI's theory of change. These questions have been designated as subordinate because are not reviewable by the TPP Evidence Review, because they are ancillary in the theory of change, or because they are not yet understood as confirmatory in the eyes of the program's developers. We investigate these questions to test hypothesized relationships and develop our theoretical understanding of how MARSSI works. Findings to these questions are necessarily of interest to the developers of MARSSI, but may also be generally informative to the field in terms of how a complex intervention like MARSSI influences behavior change. We will report results with interpretation to the program developers for future development and theory refinement. We may also choose to disseminate exploratory results more broadly in the form of peer-reviewed articles, conference papers, or possibly shorter briefs.

Modification to benchmark analytic approach. After receiving written feedback from and discussing options for a benchmark analytic approach with peer reviewers, we have decided to estimate impacts using ordinary least squares (OLS) regression for each of our three primary outcomes. We will use count models as sensitivity tests to test the robustness of those estimates. Any substantive differences between benchmark and sensitivity estimates will be reported in the results section of our impact manuscript.

Correction in the description of blocking variables. In reviewing our analysis plan, we noticed that we had mischaracterized how we intended to construct variables to account for the stratification of random assignment procedures in our analytic models. Prior to random allocation, participants were stratified sequentially into separate pools by: the state in which they indicated they would seek reproductive health care (51 possible options with the inclusion of Washington, DC) and the study coordinator who enrolled them into the study (5 options). Operationally, this has created separate randomization lists for each state-coordinator combination, for a maximum

potential set of 255 separate randomization blocks (51 states * 5 coordinators). In revising this plan, we have clarified that the dummy blocking variables we intend to include in our analytic models reflect this randomization structure.

Rationale for assessing motivation as a short-term outcome. In our list of secondary research questions, we indicate that we will assess three outcomes - motivation to use prescription birth control, motivation to use condoms, and motivation to abstain from sex – at just one time point (three months post-intervention). Our rationale for limiting this assessment to the single time point is based on two factors. First, in conversations with the intervention developer on the MARSSI theory of change, we discussed how motivation is a behavioral antecedent that likely takes time to be realized and isn't expected to shift immediately after receiving an intervention. MARSSI aims to first build an individual's confidence that they can use contraceptives and communicate effectively with their sexual partners. Once this confidence has been built, they may feel more motivated (over time) to use contraceptives and/or abstain from risky sex.

Second, we were concerned about both the burden of response on the participant and cognizant of limiting items to only those most critical to assess at specific time points.

Defining participation threshold for Complier Average Causal Effect (CACE) analyses. The CACE analyses we intend to conduct (and which are described below) are exploratory in nature; our intent is not for them to be used as our benchmark approach. Given this, we plan to assess the impact of varying levels of MARSSI participation (dosage) on outcomes of interest to better understand the potential dose-response relationship and do not specify a set threshold for participation.

Impact and Program Implementation Evaluation Analysis Plan for Boston Children's Hospital – June 2023

The Evaluation of the Momentary Affect Regulation Safer Sex Intervention in the United States

1. Impact Study Research Questions

This section presents the primary and secondary research questions that will be assessed in the impact evaluation of the Momentary Affect Regulation Safer Sex Intervention (MARSSI). Our classification of primary and secondary research questions aligns with categorization rules provided in the Instructions For TPP20 Tier 2 Phase II Impact And Program Implementation Evaluation Analysis Plan.[1] While the Participant Questionnaire collects data on many outcomes that may be affected by the intervention, following categorization rules, the research team has designated a select set of short-term behavioral outcomes as primary, which the program developer (in consultation with the research team) believes are most important for understanding the efficacy of MARSSI to achieve its goal of reducing unplanned pregnancy and STIs in the target population. Secondary research questions cover the broader array of outcomes measured in the Participant Questionnaire, as well as alternative analytic methods, that are designed to explore the impact of the program on other theoretically important behaviors and antecedents within the full sample and subgroups of interest.

a. Primary research questions[2]

1. What is the short-term impact (three months post-intervention) of the offer to participate in MARSSI (treatment) relative to the offer to participate in the control condition on participants' reported frequency of having vaginal sex without condoms in the past 30 days?
2. What is the short-term impact (three months post-intervention) of the offer to participate in MARSSI (treatment) relative to the offer to participate in the control condition on participants' reported current use of effective non-barrier contraception?
3. What is the short-term impact (three months post-intervention) of the offer to participate in MARSSI (treatment) relative to the offer to participate in the control condition on participants' reported number of sexual partners in the past 3 months?

[1] According to guidance for *Impact study research questions* (p3), in order to reduce opportunities for Type I errors and simplify reporting, researchers should designate as primary, research questions that are "focused on the outcomes most important to gauging a program's effectiveness in improving adolescent reproductive health." Findings for primary research questions are the basis for assessing program efficacy and "will be used to guide interpretation and conclusions about the effectiveness of the program being tested." By contrast, "secondary research questions examine impacts on other outcomes (aside from those examined as primary research questions) that might be influenced by the intervention or other justifiable explorations of program effectiveness." Guidance (p4) suggests that secondary outcomes are not considered "critical" to evaluating effectiveness but are none the less important to key interested parties (e.g., grantee or researchers).

[2] The primary outcomes are the principal measures on which we will assess the impact of the MARSSI intervention. These reflect sexual and reproductive health (SRH) behaviors targeted by MARSSI and identified in the logic model as the causal mechanisms that explain longer term SRH program objectives. Primary outcomes have been pre-specified in our design summary and the study's clinicaltrials.gov registration (NCT04798248).

b. Secondary research questions[3]

We include below three groups of secondary research questions, which will be used to assess the impact of MARSSI on different outcomes at different time points and with different groups. The first group of secondary research questions will explore the impact of MARSSI on the full analytic sample for specified outcomes using an intent-to-treat (ITT) framework. The second group of questions will explore the impact of MARSSI on specified subgroups, again using an ITT framework. The third group of questions will explore the impact of MARSSI using a complier average causal effect (CACE) analytic approach.

Outcomes that are starred (*) indicate those for which the construction of the measure will result in the formation of an endogenous subgroup because inclusion in the group will be defined by a participant's response to items at follow-up time points. We intend to use principal stratification analysis to estimate program effects for these measures.

Analysis of secondary outcomes in full sample using ITT approach

1. What are the post-intervention impacts of the offer to participate in MARSSI (treatment) relative to the offer to participate in the control condition on the following theoretical antecedents of sexual health and health behavior?
    a. Sexual communication self-efficacy
    b. Condom planning self-efficacy
    c. Contraceptive planning self-efficacy
    d. Coping self-efficacy
    e. Condom knowledge
    f. Contraceptive knowledge
    g. Pregnancy ambivalence*

2. What are the short-term (three months post-intervention period) impacts of the offer to participate in MARSSI (treatment) relative to the offer to participate in the control condition on the following sexual health behaviors, health outcomes, and theoretical antecedents of sexual health and health behavior?
    a. Motivation to use prescription birth control
    b. Motivation to use condoms
    c. Motivation to purposefully abstain from sex
    d. Depressive symptoms (in past 2 weeks)
    e. Frequency of vaginal sex (in past 30 days)
    f. Frequency of vaginal sex without effective non-barrier contraceptive use (in past 30 days)
    g. Frequency of emergency contraception use after vaginal sex (in past 30 days)
    h. Frequency of dual contraception (combined condom and effective non-barrier contraceptive use) during vaginal sex (in past 30 days)*

[3] Secondary outcomes reflect: a) theoretically relevant determinants/mediators of behavior and health targeted by the program as a means of achieving both short and long-term SRH outcomes; b) additional short-term outcomes the research team hypothesizes may be affected by the program but are not considered critical in evaluating the program's efficacy to achieve its ultimate goal; c) long- term outcomes the research team hypothesizes may be affected by the program but are not as likely to be realized within the timeframe of the study; and d) alternative conceptualizations of primary outcomes. Secondary research questions are intended to more broadly explore the impacts of MARSSI to better understand how and for whom the program is working. These outcomes are not considered by the program developer or researchers to be *critical* to evaluating the efficacy of MARSSI on adolescent SRH, but they are of interest to both.

i. Use of effective non-barrier contraception during last vaginal sex
j. Use of condoms during last vaginal sex
k. No effective contraception used during last vaginal sex
l. Use of dual contraception (combined condom and effective non-barrier contraceptive use) during last vaginal sex
m. Frequency of oral sex (in past 30 days)
n. Frequency of anal sex (in past 30 days)
o. Frequency of anal sex without condoms (in past 30 days)
p. Frequency of using alcohol or drugs before having any type of sex (vaginal, anal, or oral) (in past 30 days)
q. Pregnancy ambivalence*

3. What are the long-term (six months post-intervention period) impacts of the offer to participate in MARSSI (treatment) relative to the offer to participate in the control condition on the following sexual health behaviors, health outcomes, and theoretical antecedents of sexual health and health behavior?
    a. Depressive symptoms
    b. Frequency of vaginal sex (in past 30 days)
    c. Frequency of vaginal sex without effective non-barrier contraceptive use (in past 30 days)
    d. Frequency of emergency contraception use after vaginal sex (in past 30 days)
    e. Frequency of dual contraception (combined condom and effective non-barrier contraceptive use) during vaginal sex (in past 30 days)*
    f. Use of effective non-barrier contraception during last vaginal sex
    g. Use of condoms during last vaginal sex
    h. No effective contraception used during last vaginal sex
    i. Use of dual contraception (combined condom and effective non-barrier contraceptive use) during last vaginal sex
    j. Frequency of oral sex (in past 30 days)
    k. Frequency of anal sex (in past 30 days)
    l. Frequency of anal sex without condoms (in past 30 days)
    m. Frequency of using alcohol or drugs before having any type of sex (vaginal, anal, or oral) (in past 30 days)
    n. Current use of effective non-barrier contraception
    o. Frequency of vaginal sex without condoms (in past 30 days)
    p. Number of sexual partners (in past 3 months)

Analysis of primary and secondary outcomes in subgroups using ITT approach
What are the post-program, short-term, and long-term impacts of the offer to participate in MARSSI (treatment) relative to the offer to participate in the control condition on primary and secondary outcomes for subgroups defined by the following baseline data: 1) PHQ-8

depressive symptoms score; 2) age; 3) race; 4) ethnicity; 5) gender; 6) Medicaid eligibility (proxy measure of healthcare access); 7) educational level; 8) recent receipt of sexual and reproductive health (SRH) information; 9) recent receipt of health services (mental health, SRH, and/or primary health care); 10) non-barrier contraceptive use; 11) main reason for using non-barrier contraception; 12) relationship status; 13) pregnancy intentions; and 14) feelings toward pregnancy.

CACE analysis of primary and secondary outcomes
What are the post-program, short-term, and long-term impacts of threshold participation in MARSSI on primary and secondary outcomes for a subgroup of participants assigned to treatment relative to a similar group of assigned to control?

2. Impact Study Design

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

The BEhavior And Mindfulness Health (BEAM) study is an individual randomized controlled trial in which evaluators randomly assign eligible, consenting participants to intervention or comparison conditions at a one-to-one ratio. Random assignment occurs after participants' consent and before the provision of any programming or collection of baseline data. Following submission of the baseline questionnaire, participants are implicitly informed of assignment through the discussion of next steps. Participants are enrolled on a rolling basis and randomized by the state where they would most likely seek reproductive health care and by the research coordinator who is enrolling them.[4] See the *Random assignment process* section below for further details.

The treatment condition, MARSSI, is a motivational interviewing-based intervention designed specifically for adolescent and young adult (AYA) biologic females with depressive symptoms and sexual risk behavior. It aims to enhance motivation to change risky behaviors, provide skills to address depression's effects on behavior, and prompt and reinforce health-related affect regulation, cognitive behavioral skill use, and behavior change in daily life. MARSSI has three components: 1) a manualized counseling session with a sexual/reproductive health (SRH) counselor; 2) a mobile health application (mHealth app); and 3) a booster counseling session with the SRH counselor.

The main counseling session lasts ~60 minutes and uses motivational interviewing techniques to help the AYA identify a risk-reducing goal for their sexual behavior and develop a change plan. The session also provides depression education and skills, based on cognitive behavioral therapy. Counseling sessions are conducted by SRH counselors who have at least a bachelor's degree and one to two years of experience working with young people in SRH and/or mental health settings. Counselors receive approximately 24 hours of training in the intervention through self-study (assigned readings, videos), live sessions with trainers (brief didactic

[4] From June 2021 until December 2022, individuals were randomized by the state where they would most likely seek reproductive health care *at a Planned Parenthood health center,* but this stipulation was removed from the eligibility screening question in December 2022 in order to expand recruitment opportunities. From January 2023 forward, individuals were randomized by the state where they would most likely seek reproductive health care, regardless of which type of health center.

presentations, discussions, role plays), and practice (with other counselors using assigned participant scenarios, and with a mock patient followed by an evaluation/de-brief and coaching session with trainers).

Participants download the mHealth app onto their personal smartphone during the main session; the app is intended to be used four times a day for four weeks. There are three app-prompted reports and one scheduled report per day. The reports ask the participant to report their affective states and SRH risk precursors (e.g., self-efficacy to use condoms, to use contraceptives, and to refuse sex, desire for sex and reasons). When the participant reports poor affect, low contraceptive or condom self-efficacy, pregnancy desire, or desire for sex to regulate affect, they receive automated personalized messages prompting healthy behaviors and cognitive behavioral skill use. Intended dosage is approximately two minutes per day.

The booster session is designed to last ~20 minutes and is intended to be delivered after four weeks of app use. Participants speak with the SRH counselor to review behavior and relationships, discuss progress toward the participant goal, and learn a new skill (affirmations).

As originally developed, the main counseling session was intended to be delivered in person, while the booster session was designed to be delivered either in person or by secure phone or video call. However, in response to the COVID-19 pandemic, both counseling sessions have been adapted for virtual implementation via secure video call.

The control condition is a 20-minute podcast episode from the Susan G. Komen Foundation Real Pink Podcast originally aired in June 2019. It provides introductory information regarding breast health, family health history, and the way in which family health history influences individual risk of developing breast cancer; it contains no sexual health content. Participants receive a link to the podcast via email and self-administer the intervention on their own personal devices.

The study is being conducted virtually and recruiting participants throughout the United States. To support recruitment, the study team has partnered with several Planned Parenthood (PP) affiliates. Individuals have the potential to learn about the study through both active and passive recruitment methods. Active recruitment methods include having PP staff at affiliate health centers mention the study to patients and sending text messages and patient portal messages directly to patients in PP affiliate health systems which have partnered with the study. Passive recruitment methods include hanging posters and distributing handbills at affiliate health centers and posting information on the PP website notices and study-specific or -affiliated social media accounts. All study advertisements contain a QR code and/or link to the online self-screener.

When an individual takes the online self-screener, they are automatically notified whether they are eligible for the study. If they are eligible, they are routed to a page where they enter and submit their contact information. When they submit their contact information, it is immediately available to all SRH counselors at select PP affiliates who are working on the study (study coordinators). Study coordinators self-assign themselves to an eligible individual and then use the contact information provided to reach out and schedule enrollment.

To be eligible for enrollment, individuals must meet all of the following criteria: 1) biologically able to become pregnant; 2) 16-21 years of age;[5] 3) fluent in English; 4) own a smartphone; 5) have the technical capacity to participate; 6) not currently pregnant; 7) not trying to become pregnant; 8) not have given birth during the six months preceding eligibility screening; 9) not be married or engaged to be married at the time of eligibility screening; 10) report having penile- vaginal sex in the past 3 months; 11) report having penile-vaginal sex at least once a week, on average; 12) score at least 8 on the PHQ-8 depression screening tool;[6] 13) self-report visiting a reproductive health provider in the past two years;[7] 14) report having done at least one of the following during the 3 months preceding eligibility screening: a) not used a condom every time they had sexual intercourse; b) used condoms, a diaphragm, cervical cap, spermicide, sponge, fertility awareness, or withdrawal as a primary form of birth control; c) had sexual intercourse with more than one person; d) had sexual intercourse within two hours after using drugs or alcohol; or e) been treated for a sexually transmitted infection/disease; 15) consent to participate in the study; and 16) not identified as fraudulent.[8]

The intent-to-treat sample (ITT) is comprised of eligible individuals who are enrolled into the study during the enrollment period (June 2021 to April 2023).

a. Random assignment process

i. **Unit of randomization:** Random assignment occurs at the individual participant level.

ii. **Random assignment procedure:** The Policy & Research Group (PRG) is responsible for all aspects of random assignment. Random assignment blocks of varying sizes assign participants to treatment or control condition at an equal (i.e., 1:1) assignment ratio. PRG produced this allocation list with an existing algorithm available in Stata (random allocation command, *ralloc*) and did so separately for each of the potential 255 study coordinator-state combinations.[9] The allocation lists were produced by a PRG senior research analyst and stored on a secure server.

Electronic random assignment is conducted just prior to the administration of the online baseline questionnaire and is carried out by the study coordinators. Study coordinators were each given a sequential list of unique study IDs for every U.S. state. They assign a number (based on the state where the individual would seek

[5] From June 2021 until November 2022, only individuals aged 17 to 20 years were eligible for enrollment into the study. However, in December 2022, the study team expanded the criteria to include 16 and 21 years old in an effort to recruit more participants.
[6] From June 2021 until August 2021, only individuals with a PHQ-8 score of 10 or higher were eligible for enrollment into the study. However, in September 2021, the study team expanded the criteria to include those with a score of 8 or higher in an effort to recruit more participants.
[7] From June 2021 until October 2021, only individuals who reported having a virtual or an in-person visit in the past year at a Planned Parenthood health center in Alaska, Hawaii, Idaho, Indiana, Kentucky, Washington, or Wisconsin were eligible for enrollment into the study. In November 2022, the study team expanded the criteria to include anyone who reported having a visit in the past two years at a PP health center in one of these states. The criteria were expanded again in December 2021 to include anyone who reported having a visit in any U.S. state where PP health centers are located. The criteria were expanded one final time in December 2022 to include anyone who reported receiving a virtual or in-person visit in the past two years with a reproductive health provider. All of these changes were made in an effort to recruit more participants.
[8] Review of potential fraudulence is done automatically using a feature called RelevantID ®, which is integrated in the Qualtrics online survey software that the study team uses for data collection. This was added as an eligibility criterion in December 2022, when we expanded recruitment to individuals who visited any type of reproductive health provider, in an effort to ensure that only truly eligible individuals are enrolled into the study.
[9] Blocking is done at both the coordinator- and state-level. Coordinator-level blocking is used for administrative purposes. State-level blocking is used to account for regional differences in the availability of SRH services and population characteristics.

reproductive health care) to each eligible and consented individual who is able to take the questionnaire through an online data collection platform. An individual is considered enrolled in the study and part of the ITT sample when they have been given a study ID number and completed the baseline questionnaire. Random assignment occurs when the unique study ID number is entered into a field in a web-based *Randomization Generator*. It is at this point that the ID number is associated with an assignment condition (treatment or control) in the random allocation sequence. At each baseline administration (after eligibility is confirmed), the study coordinator types in the ID number that was next on the list (going in ascending numerical order) into the study ID field of the *Randomization Generator*; the application then provides a message to the study coordinator that indicates the condition associated with that ID. While the participant completes the baseline questionnaire, the study coordinator records the participant's ID number and allocation into the study's electronic *Consent and Enrollment Form*.

Study participants are not explicitly informed of which condition (treatment or control) they are assigned; however, following their submission of the baseline questionnaire, a study coordinator implicitly informs the participant of assignment through the discussion of next steps (e.g., treatment participants are provided with the MARSSI main counseling session, control participants are provided with information for how they can listen to the podcast episode and are informed that they will be contacted in a month to complete the first follow-up questionnaire, to be administered one month post-baseline). With regards to messaging, the treatment and control interventions are never discussed as such; instead, study coordinators are trained to acknowledge that participants are randomly assigned to a particular intervention: a health program involving one-on-one video sessions with the study coordinator; or a health program that involves listening to a 20- minute podcast episode.

In addition to carrying out the random assignment, PP-based study coordinators are responsible for monitoring the assignment process and ensuring fidelity to assignment is maintained. Every two weeks throughout the enrollment period, a PRG senior research analyst reviews participant IDs and assignment allocations recorded by study coordinators in the *Consent and Enrollment Form* database to ensure that they are both consistent with those in the original assignment list, and that participant IDs are assigned in sequential order as intended.

iii. **Probability of assignment to treatment group**: The probability of assignment to the treatment group is intended to be equal to the probability of assignment to the control condition; that is, p (assignment to treatment) =.5.

iv. **Potential for crossover/contamination:** To mitigate potential for both crossover[10] and contamination[11], PRG study coordinators are trained in detailed study procedures;

[10] PRG monitors crossover through routine data monitoring. Every two weeks, a senior research analyst checks to ensure that individuals assigned to the treatment condition are offered the MARSSI intervention, and that those assigned to the control condition are offered the podcast episode.

[11] In addition to crossover, it is possible that contamination may occur through interaction among participants. While it is possible

expectations regarding randomization are clear before implementation. Study coordinators commit to ensuring that participants randomly assigned to treatment participate in MARSSI and receive the intervention to fidelity, and participants randomly assigned to control receive information about how to access the podcast episode and do not receive any MARSSI content. Despite these efforts, there still remains the possibility that MARSSI participants may not receive some or all of the intervention program components, and control participants may receive some MARSSI content if study procedures are not followed by study coordinators or if SRH similar to MARSSI content is inadvertently delivered to a control participant.

that some participants may know each other, the evaluators believe that the private and personal nature of the intervention coupled with the personalization of messaging received in one-on-one sessions serves to deter the diffusion of MARSSI content from treatment to control participants. However, we do not have data that allow us to assess this issue and recognize it is plausible that treatment may be diffused in some way.

3. Program Implementation Analysis

This section lays out plans for analyzing implementation data for understanding and documenting program implementation.

Table 1. Planned implementation analysis

Research question	Measure	Operationalization	Data sources
Fidelity			
How much of the program was delivered as intended?	Counselor- and observer-reported average percent of required MARSSI activities that are completed, by session and overall	For each individual session, the proportion of required activities completed will be calculated as the number of required activities that the counselor/observer reports were completed divided by the total number of required activities for that session. There are 13 total activities that can be completed in the main counseling session and 10 total activities that can be completed in the booster session. Activities are counted as having been 'completed' if the counselor/observer reports that the activity has been completed in full. The average percent of required MARSSI activities completed will be calculated for each session type (main counseling or booster session) and overall (combining main counseling and booster session).	Counselors complete an electronic *MARSSI Main Counseling Session Self-Report Form* after delivering a main counseling session and an electronic *MARSSI Booster Session Self-Report Form* after delivering a booster session.[12] Observers complete an electronic *MARSSI Main Counseling Session Observer Form* after observing a selected main counseling session and an electronic *MARSSI Booster Session Observer Form* after observing a selected booster session.[13]
	Percent of MARSSI participants who received their booster	The number of treatment group participants who completed their booster session, divided by the total number of treatment group participants	• *Counselor Self-Report Form – Booster Session Form*

[12] Note: For the purposes of this study, delivery as intended requires that participants complete all MARSSI program components within six weeks of being enrolled into the study.

[13] Session observation entails trained MARSSI observers watching recordings of the counselors delivering sessions to participants (virtually). Counselors are asked to record a certain number of each type of session per month, with the goal of recording and observing a minimum of 10% of all main counseling sessions and booster sessions delivered during the study period. Each recording is first viewed by two observers and then data are reconciled so that there is one final Observation Form associated with each video recording. Observers complete Observer Forms within one month after each recording.

Research question	Measure	Operationalization	Data sources
	counseling session during designated two-week window		
Dosage			
How much programming was received by participants?	Percent of participants who completed the main counseling session	Number of treatment participants who are indicated as having received the main counseling session divided by the total number of treatment participants enrolled in the study	Counselors complete an electronic *Consent & Enrollment Session Form* in which they indicate if they delivered the main counseling session and an electronic *One-Month Follow-Up Data Collection Form* in which they indicate if they delivered the booster session. App usage dated are collected in the MetricWire platform and a dataset with select variables is sent to PRG on a monthly basis.
	Percent of participants who completed at least 14 days of app use	Number of treatment participants who have an indicator of at least 14 days of submitting at least one response in the app divided by the total number of treatment participants enrolled in the study[14]	
	Percent of participants who completed the booster session	Number of treatment participants who are indicated as having the received booster session divided by the total number of treatment participants enrolled in the study	
	Average percent of all sessions attended by treatment participants	For each treatment participant: • Sum of main session completed (counted as 2 sessions)[15] + indicator of at least 14 days of submitting at least one response in the app + booster session completed = total sessions completed • Total sessions completed/4 total sessions possible = Proportion of sessions completed by each participant Overall average percentage will be calculated for treatment group participants.	

[14] Although the app is set to deliver three app-prompted reports and one scheduled report per day, the intervention developer has indicated that sufficient app use is considered to be at least one response per day to any of the reports on at least 14 days of the 28-day app use period.

[15] This main counseling session is counted as 2 sessions for the two reasons. First, the intended length of the main session is 60 minutes, which is two times the intended length of the booster session (30 minutes). Secondly, the intervention developer has indicated that the main counseling session is the most critical component of the intervention and should be weighted two times more heavily than the booster session.

Research question	Measure	Operationalization	Data sources
	Percent of participants who received at least 75% of the program	The number of treatment group participants who completed either: 1) the main counseling session and 14 days of submitting at least one response in the MARSSI app; or 2) the main session and the booster session, divided by the total number of treatment group participants.	
Quality			
What was the quality of staff–participant interactions?	Counselor- and observer-reported average quality of motivational interviewing and cognitive behavioral therapy (MI/CBT) skills and techniques used during the session, by session and overall	For each individual session, counselors/observers are asked to report on a five-point scale ranging from *not at all* (=1) to *a great extent* (=5) the degree to which the following twelve MI/CBT skills and techniques were used during the session: • Interpersonal effectiveness • Setting agenda • Pacing the session • Eliciting participant's own perspectives • Reflective listening • Use of summaries to bring together what participant said • Obtaining permission from participant to offer advice • Sensitivity to participant's concerns and understanding • Partnership with participant • Support and encouragement for participant autonomy • Affirmation of participant's strengths or efforts • Guiding participant in developing change plan Each MARSSI session delivered will receive a *Counselor-Reported MI/CBT Skills and Techniques Quality* rating and each MARSSI session observed will receive an *Observer-Reported MI/CBT Skills and Techniques Quality*	Counselors complete an electronic *MARSSI Main Counseling Session Self-Report Form* after delivering a main counseling session and an electronic *MARSSI Booster Session Self-Report Form* after delivering a booster session. Observers complete an electronic *MARSSI Main Counseling Session Observer Form* after observing a selected main counseling session and an electronic *MARSSI Booster Session Observer Form* after observing a selected booster session.

Research question	Measure	Operationalization	Data sources
		rating. Both of these ratings are calculated as the average score of all twelve items in the form. The average counselor- and observer-reported ratings will be calculated for each session type (main counseling or booster session) and overall (combining main counseling and booster session).	
	Observer-reported average quality of session implementation and delivery of information, by session and overall	For each individual session that is observed, observers are asked to indicate the quality of the session on a five-point scale, from the lowest quality (=1) to the highest quality (=5), for each of the following domains: • Clarity of implementer's explanation of activities (1=not clear, 3=somewhat clear, 5=very clear) • Extent to which implementer kept track of time during the session and activities (1=not on time, 3=some loss of time, 5=well on time) • Extent to which presentation of materials seemed rushed or hurried (1=very rushed, 3=somewhat rushed, 5=not rushed) • Extent to which participant appears to understand the material (1=little understanding, 3=some understanding, 5=good understanding) • Degree to which participant is engaged in discussion and activities (1=little participation, 3=some participation, 5=active participation) The implementer's skills are also rated on a five-point scale from *poor* (=1) to *excellent* (=5) in the following areas: ○ Knowledge of the program ○ Level of enthusiasm ○ Poise and confidence	Observers complete an electronic *MARSSI Main Counseling Session Observer Form* after observing a selected main counseling session and an electronic *MARSSI Booster Session Observer Form* after observing a selected booster session.

Research question	Measure	Operationalization	Data sources
		○ Rapport and communication ○ Effectively addresses questions and concerns A final question asks observers to report on a five-point scale from *poor* (=1) to *excellent* (=5) the overall quality of the session. Each observed MARSSI session will receive an *Observer-Reported Implementation Quality rating*, which is calculated as the average score for all of these items. (Note: implementer's skills will be an average score of all five skill areas.) The average observer-reported rating will be calculated for each session type (main counseling or booster session) and overall (combining main counseling and booster session). The rating will be reported on a scale from 1 to 5, where 1 = low quality, 3 = average quality, and 5 = high quality, based upon the following scale: • 1 = average score of 1 to 1.4 • 2 = average score of 1.5 to 2.4 • 3 = average score of 2.5 to 3.4 • 4 = average score of 3.5 to 4.4 • 5 = average score of 4.5 to 5	
	Average participant-reported quality of interaction with counselor, by session and overall	Participants are asked to indicate their level of agreement on a five-point scale, from *strongly disagree* (=1) to *strongly agree* (=5), with each of the following statements about their counselor: • The counselor treated me with respect. • I felt I could trust the counselor. • I felt that the counselor listened to what I had to say.	Participants complete an electronic *MARSSI Main Counseling Session Feedback Form* after receiving the main counseling session and an electronic *MARSSI Booster Session Feedback Form* at the

Research question	Measure	Operationalization	Data sources
		• I felt the counselor understood me. • I felt comfortable with the counselor. • I felt free to ask the counselor questions. • The counselor helped me to believe that I could change and improve my life. (*Main Counseling Session*); I felt the counselor remembered what I said in the first session. (*Booster Counseling Session*) Each session for which we have a completed feedback form will receive a *Quality of Interaction with Counselor rating*, which is calculated as the average score for all seven items. The average participant-reported rating will be calculated for each session type (main counseling or booster session) and overall (combining main counseling and booster session). The rating will be reported on a scale from 1 to 5, where 1 = low quality, 3 = average quality, and 5 = high quality, based on the following scale: • 1 = average score of 1 to 1.4 • 2 = average score of 1.5 to 2.4 • 3 = average score of 2.5 to 3.4 • 4 = average score of 3.5 to 4.4 • 5 = average score of 4.5 to 5	end of the booster session delivery window.
What was quality of participant engagement with program?	Average participant-reported engagement with the main counseling session	Participants are asked to indicate their level of agreement on a five-point scale, from *strongly disagree* (=1) to *strongly agree* (=5), for each of the following statements about the MARSSI main counseling session activities: • Our discussion helped me understand ways to change my behavior.	Participants complete an electronic *MARSSI Main Counseling Session Feedback Form* after receiving the main counseling session.

Research question	Measure	Operationalization	Data sources
		• I could use what the counselor and I discussed in my daily life. • The discussion about effective contraception methods was helpful.* • The discussion about using condoms correctly and consistently was helpful.* • The discussion about choosing not to have sex was helpful.* • The discussion about having healthy relationships was helpful.* • The discussion about talking about sex with my partner was helpful.* • My Game Plan helped me think about how I will change the behavior I chose to change.* • The audio exercise(s) helped me to relax during the discussion.* • Learning the Catch It, Check It, Change It! skill for managing unhelpful thoughts was helpful.* • Role-playing a conversation with the counselor acting as my partner was helpful.* • Overall, the session with the counselor was helpful. Each participant for whom we have a completed feedback form will receive a *Participant Engagement with the Main Counseling session rating*. This will be calculated as the average score for all items in which the participant did not select the response "Not covered in my session"; items marked with a * are items where this is an available response. The rating will be reported on a scale from 1 to 5, where 1 = low engagement, 3 = average	

Research question	Measure	Operationalization	Data sources
		engagement, and 5 = high engagement, based on the following scale: • 1 = average score of 1 to 1.4 • 2 = average score of 1.5 to 2.4 • 3 = average score of 2.5 to 3.4 • 4 = average score of 3.5 to 4.4 • 5 = average score of 4.5 to 5	
	Average participant-reported engagement with MARSSI smartphone app	Participants are asked to indicate their level of agreement on a five-point scale, from *strongly disagree* (=1) to *strongly agree* (=5), for each of the following statements about the MARSSI smartphone app: • I liked completing the app phone surveys. • The app phone surveys were too long. (reverse-coded) • The app phone surveys helped me to become more aware of my thoughts, feelings, and behaviors. • There were too many app phone surveys. (reverse-coded) • I was annoyed by the app notifications. (reverse-coded) • The app messages helped me to feel better. • The app messages helped me to make the changes I want to make. • There were too many app messages. (reverse-coded) • I liked the way the app messages were written. Each participant for whom we have a completed feedback form will receive a *Participant Engagement with the App rating*. The rating will be reported on a scale from 1 to 5, where 1 = low engagement, 3 = average	Participants complete an electronic *MARSSI Booster Session Feedback Form* at the end of the booster session delivery window.

Research question	Measure	Operationalization	Data sources
		engagement, and 5 = high engagement, based on the following scale: • 1 = average score of 1 to 1.4 • 2 = average score of 1.5 to 2.4 • 3 = average score of 2.5 to 3.4 • 4 = average score of 3.5 to 4.4 • 5 = average score of 4.5 to 5	
Contrast and Context			
What other sexual and/or reproductive health programming was available or offered to study participants?	Percent of participants reporting exposure to sexual/reproductive health topics in the recent past, reported by data collection time point (baseline, four-, and seven-month follow-up)	Participants are asked whether they received the following sexual/reproductive health topics in the: 1) past 12 months at baseline; 2) past 3 months at the four months post-baseline time point; and 3) past 6 months at the seven months post-baseline time point: • Abstinence from sex or how to avoid having sex • Methods of birth control or where to get birth control • Condoms • Sexually transmitted diseases or infections • How to talk to a partner about consent and whether or not to have sex • How to talk to a partner about whether or not to use condoms or birth control • How to say no to sex • Pregnancy • Safe sexual relationships The percent of participants who report receiving each of these topics will be calculated as the number reporting exposure to the topic at a particular time point divided by the number of participants who completed the questionnaire at that respective time point.	Participants complete a *Participant Questionnaire* at baseline, one month post-baseline, four months post-baseline, and seven months post-baseline

Research question	Measure	Operationalization	Data sources
	Percent of participants reporting receipt of health services in the recent past, reported by data collection time point (baseline, four-, and seven-month follow-up)	Participants are asked whether they received the following health services in the: 1) past 12 months at baseline; 2) past 3 months at the four months post-baseline time point; and 3) past 6 months at the seven months post-baseline time point: • Mental health services • Sexual/reproductive health services • Primary health care services The percent of participants who report receiving these services will be calculated as the number reporting receipt of a certain service at a particular time point divided by the number of participants who completed the questionnaire at that respective time point.	
What external events affected implementation?	List of external events that may have affected program implementation	Description of external events that occurred during the project period that could have affected participants' receipt of programming or reproductive health services.	The PRG Lead Research Analyst records any notable external events that may have affected program implementation in a *Project Log.* PRG Research Assistants also log policies and legislation that may affect individual access to contraception and abortion services.

4. Impact Analysis

This section lays out plans for cleaning data and handling missing data, constructing outcomes, defining the analytic sample, assessing baseline equivalence, addressing potential crossover and contamination, and finally the analytic models for estimating program impacts and planned sensitivity analyses.

a. Data cleaning

Prior to analysis, PRG staff will systematically screen or review the analytic data (baseline and outcome) to identify invalid, outlying, missing, and unreliable observations.[16] In our benchmark approach, new variables are created in which data that are deemed unusable (i.e., invalid) are coded as missing and flagged according to missing data type; all other data are retained, unchanged. Missing data that are not invalid will be updated through logical editing when possible, and all missing covariate data will be updated using dummy variable adjustments. Details of our data cleaning steps and rationale for our missing data approach are outlined below. In addition, we will assess the robustness of these analytic decisions with sensitivity analyses and report on any substantive inconsistencies, as detailed in the *Sensitivity analyses* section.

i. **Item-Level Procedures.** Data cleaning begins with a thorough review of all questionnaire items. The goal of item-level procedures is to prepare data for analytic variable construction. To this end, we ensure data are as complete as possible and that all recorded values are valid.

   a. **Identify and flag invalid responses**: The first step in the data screening process is to inspect the data for instances in which responses are invalid because they are outside of a pre-determined range of plausible or acceptable values. Each questionnaire type (e.g., baseline, follow-up) has a codebook, prepared by a PRG staff person, that contains variable names, pre-specified and valid variable values or ranges of values, and when applicable, value labels.[17] Referring to the codebook, a research analyst performs diagnostics in Stata to ensure that values for all variables used in analysis are valid (i.e., data are within ranges specified in the codebook). A data analyst inspects the data using two commands in Stata. First, the analyst uses the command *sum variable_name,* which provides summary statistics (mean, minimum, maximum, standard deviation) for all numeric variables. The analyst checks that the minimum and maximum values are valid. If this command reveals there are values out of range, the analyst then inspects the

[16] With regards to the potential for inconsistent responses, during instrument construction, the study team considered what types of questions may lead to inconsistencies – both internal (within the same instrument) and over-time (across instrument) inconsistencies. To avoid internal inconsistencies in our primary outcomes, we built skip patterns into the online questionnaire. If participants indicate they have not recently had a particular type of sex, they are skipped out of more specific questions related to that type of sex. If they state they have not recently had any sexual partners, they are then skipped out of a question asking about the types of sexual partners they have had. In addition, participants are precluded from indicating they had a particular type of sex without a condom more times than they said they had that type of sex. Furthermore, given the eligibility criteria for this study require participants to be recently sexually active, there was not a need to ask questions about whether participants have ever had certain types of sex; as a result, there are no items in the instrument that offer the potential for over-time inconsistencies to exist.

[17] Regardless as to whether data are nominal, ordinal, or continuous, all response options are coded in Stata as numeric values; values are labeled according to corresponding category names when data are nominal or ordinal. As an example, the variable gender is a nominal variable; however, each response option (female, male, transgender female; transgender male; non-binary/gendergueer, unsure/questioning, other) is coded as a pre-specified, unique numeric value (1-5). The only acceptable values for this variable then are 1-7; any other values are out of range.

data using the command, *tab variable_name, missing*, which provides a frequency table of all values (including missing values) so the analyst can identify and flag all values that are out of range as invalid and recode these values to missing (code as ".k"). Data that are recoded to missing are treated according to our missing data approach. Briefly, our benchmark approach is to impute missing baseline data and include in analysis; we exclude observations with missing outcome data from analysis.

b. **Assess missingness:** The second step in the data screening process is to assess missingness. In this step, a research analyst reviews and reports the prevalence of unit and item missingness (which results from nonresponse) for both treatment and control samples.

c. **Conduct logical data edits:** The third step is to determine if logical edits are possible for any variables that may have missing values due to skip patterns and nonresponse and logically edit where that may be the case.[18] We will not logically edit where the missing values are previously determined to be invalid.

ii. **Analytic variable level procedures.** After review and updates to individual items, we construct our analytic variables and review resulting measures for outliers.[19] Outliers are values that are extreme compared to other observations but are not plainly invalid. In the data cleaning process, we inspect outliers so that we can try to ascertain whether they are in fact true (or plausible) values or potentially a result of measurement error. The only variables for which we inspect outliers are those used in the construction of our outcome variables (frequency of having vaginal sex, frequency of having vaginal sex without condoms, frequency of having vaginal sex without being protected by some form of prescription birth control, frequency of using emergency contraception after vaginal sex, frequency of having vaginal sex using both a condom and prescription birth control at the same time, frequency of having oral sex, frequency of having anal sex, frequency of having anal sex without condoms, frequency of using drugs or alcohol before any type of sex, and number of sexual partners) because they have no upper limit (all other variables used in analysis are either categorical or have predicated upper and lower bounds). Our approach is to identify and flag influential observations in our data.

Our benchmark analytic approach is to include data flagged as outliers in analysis, because we do not know for certain whether the values are true or invalid. However, we also run sensitivity analyses that exclude these data and report substantive differences in the results section of the report.

iii. **Instrument level procedures.** Once analytic variables have been constructed and reviewed, we review entire cases to determine if they are reliable and conduct dummy variable adjustment on our baseline analytic variables.

[18] PRG's general approach to logically editing a specific variable is to use as few other variables as possible. Variables that are missing due to a skip pattern are updated to their logical value of zero. For variables in which an item should have been answered but was not, we use only the variable that directly preceded it to update its value.

[19] After constructing our analytic variables, we conduct the first step of the data review again (identify and flag invalid responses). Although this has already been done at the item-level, this additional check allows a better understanding of variable construction.

a. **Identify and flag unreliable cases:** The final step in the data screening process is to identify and flag entire cases (i.e., entire questionnaires) that are unreliable. By unreliable, we mean that we have sufficient reason to believe that the respondent's answers are not honest representations of their behaviors, knowledge, and beliefs. Cases are flagged as unreliable if the participant indicates they have not been honest in responding to the questionnaire or if project staff indicate in project logs specific issues encountered during data collection that are cause for treating a case as unreliable. Honesty during questionnaire administration is assessed from the following item on the *Participant Questionnaire*.

   - *Have you been as honest as possible in responding to all of the questions in this questionnaire?*

   Persons who indicate *No*, *none of the time* are flagged as unreliable. In addition, cases may be flagged as unreliable if project log notes indicate a notable issue during data collection (e.g., participant determined to be fraudulent). Each suspected unreliable case is reviewed by the project lead and a senior analyst for a final determination of reliability. Data for cases that are deemed unreliable are included in the benchmark analyses. However, sensitivity analyses that exclude the unreliable data will be conducted and results will be reported in an appendix of the report.

b. **Adjustment of baseline data:** In the final step of the data cleaning process, we determine if any individuals who are in the randomized sample (for each outcome) are missing baseline covariates or the baseline measure of the outcome variable. If this is the case, our proposed benchmark approach is to use dummy variable adjustment procedures, i.e., we code missing data to either zero or the mean of non-missing observations (for dichotomous and count/continuous variables, respectively). We construct dummy indicators to identify missing cases imputed to the constant/mean value and we adjust estimates by including these dummy indicators in analytic models.

iv. **Missing data approach:** Assuming that our study design and procedures are sound, missing data pose perhaps the greatest threat to the internal validity of our RCT study and the ITT framework (Puma et al. 2009; Moher et al., 2010).[20] Randomization at the point of offer allows us to make causal statements about the effect of that offer because treatment and comparison samples are equal in expectation. For the ITT framework to remain internally valid, however, the treatment and comparison groups must remain equal in expectation at the point of analysis. When the analytic sample is diminished by attrition or non-response, non-random differences (i.e., self-selecting) between the treatment and comparison groups may be introduced into the sample and estimates of program impacts may become biased. Although there is no consensus on how to resolve

[20] Puma, M.J., Olsen, R.B., Bell, S.H., Price, C. (2009). What to Do When Data Are Missing in Group Randomized Controlled Trials. (NCEE 2009-0049). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Moher, D. et al. (2010). CONSORT 2010 Explanation and Elaboration: Updated Guidelines for Reporting Parallel Group Randomised Trials. BMJ 2010;340:c869.

this, practical guidance on how to address and mitigate the problems associated with missing data have been published in education (Puma et al., 2009).

Our six-step decision process for addressing this problem, as detailed below, is informed by this guidance. These steps (which are incorporated into our data cleaning procedures) articulate how we will deal with missing outcome and baseline/covariate data (variables outlined in the *Model specification and covariates* section, necessary for the estimation of impacts). The benchmark approach that we have selected aims to mitigate the introduction of bias into our impact estimates and maximize the use of available data by adjusting missing baseline/covariate data. To test the robustness of this approach, and to verify these findings, we will report comparative findings using sensitivity analyses that also employ an alternative method which includes no adjustment (as outlined in step 6).

a. Using data cleaning procedures outlined in the *Data cleaning* section, identify outlying, unreliable, and invalid data in any analytic (i.e., outcome, baseline, or covariate) variables. Recode invalid data as missing, and flag outlier and unreliable data for sensitivity analyses.[21]

b. Report prevalence of unit and item missingness (which result from nonresponse and invalid data) for both treatment and control samples.[22]

c. Determine if logical edits are possible for any analytic variables that may have missing values (due to nonresponse) and logically update missing values where this is the case. We will not logically impute where the missing values are previously invalid.

d. Determine if any individuals who are in the randomized sample (for each outcome) do not have outcome data at the follow-up time point. If this is the case, our proposed benchmark approach is to use case deletion, as we feel it is the most straightforward and prudent approach for missing follow-up data recommended in Puma et al. (2009). These cases will be deleted from the analytic sample and attrition statistics will be reported.

e. Determine if any individuals who are in the analytic sample (for each outcome) are missing baseline covariates or the baseline measure of the outcome variable. If this is the case, our proposed benchmark approach is to use dummy variable adjustment procedures, as we feel it is the most straightforward and prudent approach for missing baseline/covariate data recommended in Puma et al. (2009).

f. Conduct sensitivity analyses by estimating results with missing baseline data excluded from the analysis (i.e., use case-wise deletion for all cases with missing baseline and outcome data). In an appendix, we will report our benchmark results next to the sensitivity analysis results to verify findings.

[21] We will code missing responses with a unique missing code that identifies or flags these missing values according to the reason they are missing (i.e., nonresponse, invalid, inconsistent). See the *Data cleaning* section for details on how missing data are coded.
[22] For item missing values, we will only report prevalence of missing data for variables that are included in our model specifications and could therefore influence the constitution of the analytic sample.

b. Outcome measures

Our primary research questions ask to what extent the offer to participate in MARSSI relative to the offer to listen to a breast health podcast impacts participants' reported: 1) frequency of having vaginal sex without a condom (in past 30 days); 2) current use of effective non-barrier contraception; and 3) number of sexual partners (in past 3 months), at short-term follow-up (three months post-intervention period). Below, we describe the specific operationalization of these three outcome measures.

Frequency of having vaginal sex without condoms (in past 30 days)

We operationalize condom use as a risk outcome; that is, we measure the frequency with which participants engage in the risk behavior of having vaginal sex without a condom, rather than the frequency with which they engage in the safe sex practice of using condoms. Constructing the variable in this way allows us to examine the self-reported sexual behaviors of the full analytic sample of participants, regardless as to whether or not they are recently sexually active.

Specifically, the frequency of having vaginal sex without condoms is constructed as a count measure - the total number of times a respondent reports *not* using a condom during vaginal sex in the past 30 days. As constructed, the measure of risk is cumulative.[23] A score of 0 indicates no risk (i.e., the individual has not engaged in vaginal sex without a condom either because they are sexually active and always use condoms or they are not sexually active); higher values indicate more discrete instances of risk. Data used to assess the impact of the treatment (MARSSI) on condom use during vaginal sex are obtained from the following item on the *Participant Questionnaire,* which is administered to both the treatment and control groups at baseline, short-term follow-up (three months post-intervention period), and long-term follow-up (six months post-intervention period).

- *In the past 30 days, how many times have you had vaginal sex without using a condom?*

Persons who indicate that they have not had vaginal sex in the past 30 days are coded as having vaginal sex without a condom zero times.[24] MARSSI will be considered to have a positive impact on frequency of having vaginal sex without condoms if the frequency (number of times) of vaginal sex without condoms reported by participants assigned to MARSSI three months post-intervention period is smaller than the frequency reported by control participants and the difference between groups is statistically significant.

Current use of effective non-barrier contraception

Use of effective non-barrier contraception is constructed as a dichotomous variable –

[23] We have elected to construct vaginal sex without a condom as a count instead of proportionate measure because we believe it provides a clearer depiction of an individual's magnitude of risk. For example, if we operationalized risk as a proportionate measure, an individual who reports having vaginal sex two times in the past 30 days, once without a condom, would have the same risk ratio as an individual who reports having vaginal sex 50 times in the past 30 days, with 25 instances when a condom wasn't used (50% risk). Operationalizing risk as a cumulative measure allows us to take into account the frequency with which an individual is having sex.

[24] The *Participant Questionnaire* contains sexual behavior questions that use a 30-day recall period. Research has consistently found that memory of behaviors/events decreases over time and accuracy of recall is negatively associated with length of recall period (Clarke et al. 2008; Schwarz and Oyserman 2001), especially for more frequent behaviors. Since participants must have reported engagement in recent vaginal sex and some sexual risk to be eligible, we assume for this group sexual activity is more common. As such, we use items with a 30-day recall period to construct our measures of sexual behaviors as we believe these should elicit more accurate responses than a potentially more traditional, but longer recall period (e.g., three-month).

participants are either coded as currently using effective non-barrier contraception or not currently using effective non-barrier contraception. Data used to assess the impact of the treatment (MARSSI) on contraceptive use are obtained from the following item on the *Participant Questionnaire*, which is administered to both the treatment and control groups at baseline, short-term follow-up (three months post-intervention period), and long-term follow-up (six months post-intervention period).

- *Please indicate which method of prescription birth control you are currently using.*
    - *Oral contraceptives (for example, the pill)*
    - *The patch (for example, Ortho Evra)*
    - *The shot/injection (for example, Depo-Provera)*
    - *The ring (for example, NuvaRing)*
    - *The implant (for example, Implanon or Nexplanon)*
    - *IUD (for example, ParaGard, Skyla, Mirena, Kyleena, Liletta)*
    - *None of the above*

Persons who indicate that they are currently using any of the listed methods are considered to be currently using effective non-barrier contraception and are coded as 1. Persons who select *None* are considered to not be currently using effective non-barrier contraception and are coded as 0.

MARSSI will be considered to have a positive impact on non-barrier contraceptive use if, as compared to participants who are assigned to the control group, a larger proportion of participants who are offered MARSSI report using effective non-barrier contraception at the three-month post-intervention period and the difference between groups is statically significant.

Number of sexual partners

Number of sexual partners is constructed as a count variable – the number of sexual partners the participant reports that they have had in the past three months.[25] Data used to assess the impact of the treatment (MARSSI) on number of sexual partners are obtained from the following item on the *Participant Questionnaire,* which is administered to both the treatment and control groups at baseline, short-term follow-up (three months post-intervention period), and long-term follow-up (six months post-intervention period).

- *How many sexual partners have you had in the past 3 months?*

Persons who indicate that they have not had any sexual partners in the past three months are coded as having zero sexual partners.

MARSSI will be considered to have a positive impact on number of sexual partners in the past three months if the number of sexual partners reported by participants assigned to MARSSI at the three-month post-intervention follow-up is smaller than the number of sexual partners reported by control participants and the difference between groups is statistically significant.

[25] We elected to use a three-month recall period for this outcome measure as it measures a behavior that is likely to occur with less frequency than individual acts of sex, and thus represents a quantity that would be easier to recall over a longer period of time.

Table 2. Outcomes used for primary research questions

Outcome name	Source item(s)	Constructed measure	Timing of measure
Frequency of having vaginal sex without condoms in the past 30 days	*Participant Questionnaire*	The risk outcome is operationalized as the number of times in the past 30 days a person reports having vaginal sex without using a condom. The measure is calculated from the following item: • In the past 30 days, how many times have you had vaginal sex without using a condom? The resulting variable is continuous with values that range from 0 to *k*, where 0 indicates that a person has not engaged in vaginal sex without a condom in the past 30 days, and *k* indicates the number of times the person has engaged in vaginal sex without a condom (risk behavior) in the past 30 days. Note: All respondents who have three-month post-intervention follow-up data and have provided a response to either this question or have indicated they have not had vaginal sex in the past 30 days will be included in the construction of this measure. Persons who indicate that they have not had vaginal sex in the past 30 days are coded as having vaginal sex without a condom zero times .	Three months post-intervention (four months post-baseline)

Current use of effective non-barrier contraception	*Participant Questionnaire*	The protective outcome is operationalized as a dichotomous variable indicating whether a person reports currently using effective non-barrier contraception or not.
		
		The measure is calculated from the following item:
		• Please indicate which method of prescription birth control you currently using:
		  ○ Oral contraceptives (for example, the pill)
		  ○ The patch (for example, Ortho Evra)
		  ○ The shot/injection (for example, Depo-Provera)
		  ○ The ring (for example, NuvaRing)
		  ○ The implant (for example, Implanon or Nexplanon)
		  ○ IUD (for example, ParaGard, Skyla, Mirena, Kyleena, Liletta)
		  ○ None of the above
		
		A person who selects *Oral contraceptives, patch, shot/injection, ring, implant,* or *IUD* is given a value of 1 for the measure. A person who selects *None* is given a value of 0 for the measure.
		
		The resulting variable is dichotomous with values 0 or 1, where 0 indicates a person who does not currently use effective non-barrier contraception and 1 indicates a person who does currently use effective non-barrier contraception.
		
		Note: All respondents who have three-month post-intervention follow-up data and have provided a response to this question will be included in the construction of this measure.
			Three months post-intervention (four months post-baseline)

Number of sexual partners in the past 3 months	*Participant Questionnaire*	The risk outcome is operationalized as the number of sexual partners in the past three months. The measure is calculated from the following item: • How many sexual partners have you had in the past 3 months? The resulting measure is the total number of sexual partners reported by the participant. The resulting variable is continuous with values that range from 0 to *k*, where 0 indicates that a person has had no sexual partners in the past three months, and *k* indicates the number of sexual partners in the past three months. Note: All respondents who have three-month post-intervention follow-up data and have provided a response to this question will be included in the construction of this measure.	Three months post-intervention (four months post-baseline)

Table 3. Outcomes used for secondary research questions[26]

Outcome name	Source item(s)	Constructed measure	Timing of measure
Sexual communication self-efficacy	*Participant Questionnaire*	Participants are asked the following series of questions and asked to rate on a 7-point scale, how confident or sure are they that they could: • Tell someone you plan to have sex with that you want to use condoms. • Convince a partner to use condoms, even if you are using some other kind of birth control (for example, the pill) • Insist that a condom be used • Refuse to have sex if a partner won't use a condom • Tell a partner that you do not want to have sex Response values range from 1 to 7, where 1 indicates that the respondent is *not at all confident*, and 7 indicates the respondent is *extremely confident*. The measure is calculated as the average response to all five items.	Post-intervention (one month post-baseline) and three months post-intervention (four months post-baseline)
Condom planning self-efficacy	*Participant Questionnaire*	Participants are asked the following series of questions and asked to rate on a 7-point scale, how confident or sure are they that they could: • Use a condom every time that you have sex. • Use a condom correctly every time you have sex • Use a condom after you have been drinking Response values range from 1 to 7, where 1 indicates that the respondent is *not at all confident*, and 7 indicates the respondent is *extremely confident*. The measure is calculated as the average response to all three items.	Post-intervention (one month post-baseline) and three months post-intervention (four months post-baseline)
Contraceptive planning self-efficacy	*Participant Questionnaire*	Participants are asked the following series of questions and asked to rate on a 7-point scale, how confident or sure are they that they could: • Use prescription birth control as directed • Resist having sex if you are not using some form of prescription birth control Response values range from 1 to 7, where 1 indicates that the respondent is *not at all confident*, and 7 indicates the respondent is *extremely confident*. The measure is calculated as the average response to both items.	Post-intervention (one month post-baseline) and three months post-intervention (four months post-baseline)
Coping self-efficacy	*Participant Questionnaire*	Participants are asked the following series of questions and asked to rate on a 7-point scale, how confident or sure are they that they could: • Make unpleasant thoughts go away • Take your mind off unpleasant thoughts • Stop yourself from being upset by unpleasant thoughts • Keep from feeling sad Response values range from 1 to 7, where 1 indicates that the respondent is *not at all confident*, and 7 indicates the respondent is *extremely confident*. The measure is calculated as the average response to all four items.	Post-intervention (one month post-baseline) and three months post-intervention (four months post-baseline)

[26] For each of the scale measures listed in this table, while we detail the items we have included in the *Participant Questionnaire* that we intend to use to construct each measure, we plan to assess the dimensionality and internal consistency of each scale before using it in analysis to ensure they are reliable measures.

Condom knowledge	*Participant Questionnaire*	Participants are provided with the following series of statements and asked to indicate whether each is *True* or *False.* Participants may also indicate they *Don't know:* • Condoms have an expiration date. • Condoms work well to prevent sexually transmitted infections. • Condoms are not as effective at preventing pregnancy as prescription birth control methods (for example, IUDs, the implant, the pill, the patch, the ring, the shot). Correct answers are coded as 1, incorrect answers are coded as 0. The measure is constructed as the number of correct responses out of total 3 items.	Post-intervention (one month post-baseline) and three months post-intervention (four months post-baseline)
Contraceptive knowledge	*Participant Questionnaire*	Participants are provided with the following series of statements and asked to indicate whether each is *True* or *False.* Participants may also indicate they *Don't know:* • Birth control pills are effective even if a woman misses taking them for two or three days in a row. • Long-acting methods like the implant or IUD cannot be removed early, even if a woman changes her mind about wanting to get pregnant. • The birth control pill, ring, and patch are just as effective as IUDs and the implant • Plan B and Ella are pills that can be taken shortly after having unprotected sex to prevent pregnancy. • Some methods of emergency contraception (such as Plan B or Ella) require a prescription. Correct answers are coded as 1, incorrect answers are coded as 0. The measure is constructed as the number of correct responses out of total 5 items.	Post-intervention (one month post-baseline) and three months post-intervention (four months post-baseline)
Motivation to use prescription birth control	*Participant Questionnaire*	Participants are asked to indicate on a 5-point scale, how motivated they are to do the following thing in the next six months: • Use prescription birth control Response values range from 1 to 5, where 1 indicates *not at all* motivated to do this in the next six months and 5 indicates *extremely* motivated to do this in the next six months. The measure is single-item measure.	Three months post-intervention (four months post-baseline)
Motivation to use condoms	*Participant Questionnaire*	Participants are asked to indicate on a 5-point scale, how motivated they are to do the following thing in the next six months: • Use condoms if you have sex Response values range from 1 to 5, where 1 indicates *not at all* motivated to do this in the next six months and 5 indicates *extremely* motivated to do this in the next six months. The measure is single-item measure.	Three months post-intervention (four months post-baseline)
Motivation to purposefully abstain from sex	*Participant Questionnaire*	Participants are asked to indicate on a 5-point scale, how motivated they are to do the following thing in the next six months: • Purposefully abstain from sex Response values range from 1 to 5, where 1 indicates *not at all* motivated to do this in the next six months and 5 indicates *extremely* motivated to do this in the next six months. The measure is single-item measure.	Three months post-intervention (four months post-baseline)

Pregnancy ambivalence	*Participant Questionnaire*	The outcome is operationalized as a dichotomous variable indicating if someone is ambivalent about pregnancy or not.
		
		The measure is calculated from the following items:
		• Thinking about your life right now, how important is it to you to avoid becoming pregnant?
		  ○ Very important
		  ○ Somewhat important
		  ○ A little important
		  ○ Not important
		• If you found out today that you were pregnant, how would you feel?
		  ○ Very upset
		  ○ A little upset
		  ○ A little pleased
		  ○ Very pleased
		
		Respondents are coded as being unambivalent (=0) about wanting to prevent a pregnancy if they indicate that it is *Very important* for them to avoid pregnancy and that they would be *Very upset* or A *little upset* by a pregnancy.
		
		Respondents are coded as being unambivalent (n=0) about wanting a pregnancy if they indicated that it was N*ot important* for them to avoid pregnancy and that they would be V*ery pleased* or *A little pleased* about discovering a pregnancy.
		
		All other respondents are coded as ambivalent (n=1). This group includes respondents who provide inconsistent or conflicting responses to the two items (e.g., V*ery important* to avoid pregnancy yet V*ery pleased* if a pregnancy occurred), those who give midscale responses for both items (e.g., *Somewhat important* to avoid pregnancy and *A little pleased* if a pregnancy were discovered).[27] A person who indicates they are currently pregnant at that time point are excluded from this analysis.
			Post-intervention (one month post-baseline) and three months post-intervention (four months post-baseline)

[27] This scale construction is based on the work of Higgins, J. A., Popkin, R. A., & Santelli, J. S. (2012). Pregnancy ambivalence and contraceptive use among young adults in the United States. *Perspectives on sexual and reproductive health*, *44*(4), 236-243.

Depressive symptoms in the past 2 weeks	*Participant Questionnaire*	Participants are asked to indicate for each of the following items how often during the past two weeks they were bothered by: • Little interest or pleasure in doing things • Feeling down, depressed, or hopeless • Trouble falling or staying asleep, or sleeping too much • Feeling tired or having little energy • Poor appetite or overeating • Feeling bad about yourself, or that you are a failure or have let yourself or your family down • Trouble concentrating on things, such as reading the newspaper or watching television • Moving or speaking so slowly that other people could have noticed. Or the opposite – being so fidgety or restless that you have been moving around a lot more than usual Response values are 0=*not at all*, 1=*several days,* 2=*more than half the days*, 3=*nearly every day.* Scores to all eight items are summed for a summative score.	Three months post-intervention (four months post-baseline) and six months post-intervention (seven months post-baseline)
Frequency of having vaginal sex in the past 30 days	*Participant Questionnaire*	The measure is operationalized as the number of times in the past 30 days a person reports having vaginal sex. The measure is calculated from the following item: • In the past 30 days, how many times have you had vaginal sex? The resulting variable is continuous with values that range from 0 to *k*, where 0 indicates that a person has not engaged in vaginal sex in the past 30 days, and *k* indicates the number of times the person has engaged in vaginal sex in the past 30 days.	Three months post-intervention (four months post-baseline) and six months post-intervention (seven months post-baseline)
Frequency of having vaginal sex without effective non-barrier contraceptive use (in past 30 days)	*Participant Questionnaire*	The measure is operationalized as the number of times in the past 30 days a person reports not using effective non-barrier contraception during vaginal sex. The measure is calculated from the following item: • In the past 30 days, how many times have you had vaginal sex without being protected by some form of prescription birth control? The resulting variable is continuous with values that range from 0 to *k*, where 0 indicates that a person has not engaged in vaginal sex without using prescription birth control in the past 30 days, and *k* indicates the number of times the person has engaged in vaginal sex without using prescription birth control in the past 30 days. Persons who indicate that they have not had vaginal sex in the past 30 days are coded as having vaginal sex without prescription birth control zero times.	Three months post-intervention (four months post-baseline) and six months post-intervention (seven months post-baseline)

Frequency of using emergency contraception after vaginal sex in the past 30 days	*Participant Questionnaire*	The measure is operationalized as the number of times in the past 30 days a person reports using emergency contraception. The measure is calculated from the following item: • In the past 30 days, how many times have you had used emergency contraception after vaginal sex to prevent pregnancy? The resulting variable is continuous with values that range from 0 to *k*, where 0 indicates that a person has not used emergency contraception in the past 30 days, and *k* indicates the number of times the person has used emergency contraception in the past 30 days. Persons who indicate that they have not had vaginal sex in the past 30 days are coded as using emergency contraception zero times.	Three months post-intervention (four months post-baseline) and six months post-intervention (seven months post-baseline)
Frequency of using dual contraception during vaginal sex in the past 30 days	*Participant Questionnaire*	The measure is operationalized as the number of times in the past 30 days a person reports using dual contraception during vaginal sex. The measure is calculated from the following item: • In the past 30 days, how many times did you have vaginal sex using both a condom and one of the list forms of prescription birth control at the same time? The resulting variable is continuous with values that range from 0 to *k*, where 0 indicates that a person has not used dual contraception during vaginal sex in the past 30 days, and *k* indicates the number of times the person has used dual contraception in the past 30 days. Persons who indicate that they have not had vaginal sex in the past 30 days are excluded from this analysis.	Three months post-intervention (four months post-baseline) and six months post-intervention (seven months post-baseline)

Use of effective non-barrier contraception during last vaginal sex	*Participant Questionnaire*	The measure is operationalized as a dichotomous variable indicating if someone used effective non-barrier contraception or not during last vaginal sex.
		
		The measure is calculated from the following item:
		• Considering the LAST time you had vaginal sex, which of the following did you use?
		  ○ Oral contraceptives (for example, the pill)
		  ○ The patch (for example, Ortho Evra)
		  ○ The shot/injection (for example, Depo-Provera)
		  ○ The ring (for example, NuvaRing)
		  ○ Implant (for example, Implanon or Nexplanon)
			Three months post-intervention (four months post-baseline) and six months post-intervention (seven months post-baseline)

		○ IUD (for example, ParaGard, Skyla, Mirena, Kyleena, Liletta) ○ Emergency contraception (ella, Plan B) ○ Condoms ○ The sponge ○ Diaphragm ○ Foam or spermicide ○ Natural family planning (rhythm method) ○ I did not use any of these A person who selects *Oral contraceptives, The patch, The shot/injection, The ring, Implant, or IUD* is given a value of 1 for the measure. A person who selects anything else is given a value of 0 for the measure.	
Use of condoms during last vaginal sex	*Participant Questionnaire*	The measure is operationalized as a dichotomous variable indicating if someone used condoms or not during last vaginal sex. The measure is calculated from the following item: • Considering the LAST time you had vaginal sex, which of the following did you use? ○ Oral contraceptives (for example, the pill) ○ The patch (for example, Ortho Evra) ○ The shot/injection (for example, Depo-Provera) ○ The ring (for example, NuvaRing) ○ Implant (for example, Implanon or Nexplanon) ○ IUD (for example, ParaGard, Skyla, Mirena, Kyleena, Liletta) ○ Emergency contraception (ella, Plan B) ○ Condoms ○ The sponge ○ Diaphragm ○ Foam or spermicide ○ Natural family planning (rhythm method) ○ I did not use any of these A person who selects *Condoms* is given a value of 1 for the measure. A person who selects anything else is given a value of 0 for the measure.	Three months post-intervention (four months post-baseline) and six months post-intervention (seven months post-baseline)
No effective contraception used during last vaginal sex	*Participant Questionnaire*	The measure is operationalized as a dichotomous variable indicating if someone did not use effective contraception during last vaginal sex. The measure is calculated from the following item: • Considering the LAST time you had vaginal sex, which of the following did you use? ○ Oral contraceptives (for example, the pill) ○ The patch (for example, Ortho Evra) ○ The shot/injection (for example, Depo-Provera) ○ The ring (for example, NuvaRing)	Three months post-intervention (four months post-baseline) and six months post-intervention (seven months post-baseline)

		○ Implant (for example, Implanon or Nexplanon) ○ IUD (for example, ParaGard, Skyla, Mirena, Kyleena, Liletta) ○ Emergency contraception (ella, Plan B) ○ Condoms ○ The sponge ○ Diaphragm ○ Foam or spermicide ○ Natural family planning (rhythm method) ○ I did not use any of these A person who selects *I did not use any of these* is given a value of 1 for the measure. A person who selects anything else is given a value of 0 for the measure.	
Use of dual contraception during last vaginal sex	*Participant Questionnaire*	The measure is operationalized as a dichotomous variable indicating if someone used both effective non-barrier contraception AND condoms or not during last vaginal sex. The measure is calculated from the following item: • Considering the LAST time you had vaginal sex, which of the following did you use? ○ Oral contraceptives (for example, the pill) ○ The patch (for example, Ortho Evra) ○ The shot/injection (for example, Depo-Provera) ○ The ring (for example, NuvaRing) ○ Implant (for example, Implanon or Nexplanon) ○ IUD (for example, ParaGard, Skyla, Mirena, Kyleena, Liletta) ○ Emergency contraception (ella, Plan B) ○ Condoms ○ The sponge ○ Diaphragm ○ Foam or spermicide ○ Natural family planning (rhythm method) ○ I did not use any of these A person who selects <u>at least one of the following:</u> *Oral contraceptives, The patch, The shot/injection, The ring, Implant, IUD* <u>and</u> *Condoms* is given a value of 1 for the measure. A person who selects anything else is given a value of 0 for the measure.	Three months post-intervention (four months post-baseline) and six months post-intervention (seven months post-baseline)
Frequency of having oral sex in the past 30 days	*Participant Questionnaire*	The measure is operationalized as the number of times in the past 30 days a person reports having oral sex. The measure is calculated from the following item: • In the past 30 days, how many times have you had oral sex?	Three months post-intervention (four months post-baseline) and six months post-intervention

		The resulting variable is continuous with values that range from 0 to *k*, where 0 indicates that a person has not had oral sex in the past 30 days, and *k* indicates the number of times the person has had oral sex in the past 30 days.	(seven months post-baseline)
Frequency of having anal sex in the past 30 days	*Participant Questionnaire*	The measure is operationalized as the number of times in the past 30 days a person reports having anal sex. The measure is calculated from the following item: • In the past 30 days, how many times have you had anal sex? The resulting variable is continuous with values that range from 0 to *k*, where 0 indicates that a person has not had anal sex in the past 30 days, and *k* indicates the number of times the person has had anal sex in the past 30 days.	Three months post-intervention (four months post-baseline) and six months post-intervention (seven months post-baseline)
Frequency of having anal sex without condoms in the past 30 days	*Participant Questionnaire*	The measure is operationalized as the number of times in the past 30 days a person reports having anal sex without using a condom. The measure is calculated from the following item: • In the past 30 days, how many times have you had anal sex without using a condom? The resulting variable is continuous with values that range from 0 to *k*, where 0 indicates that a person has not engaged in anal sex without a condom in the past 30 days, and *k* indicates the number of times the person has engaged in anal sex without a condom (risk behavior) in the past 30 days. Persons who indicate that they have not had anal sex in the past 30 days are coded as having anal sex without condoms zero times.	Three months post-intervention (four months post-baseline) and six months post-intervention (seven months post-baseline)
Frequency of using alcohol or drugs before having any type of sex in the past 30 days	*Participant Questionnaire*	The measure is operationalized as the number of times in the past 30 days a person reports using alcohol or drugs before having any type of sex in the past 30 days. The measure is calculated from the following item: • In the past 30 days, how many times did you use alcohol or drugs before having any type of sex (vaginal, oral, or anal)? The resulting variable is continuous with values that range from 0 to *k*, where 0 indicates that a person has not used alcohol or drugs before having any type of sex in the past 30 days, and *k* indicates the number of times the person has used alcohol or drugs before having any type of sex (risk behavior) in the past 30 days. Persons who indicate that they have not had vaginal, anal, and oral sex in the past 30 days are coded as having using alcohol or drugs before any type of sex zero times.	Three months post-intervention (four months post-baseline) and six months post-intervention (seven months post-baseline)

c. Analytic sample(s)

The study sample is comprised of individuals who have met the study eligibility criteria (see *Section 2. Impact Study Design* for a detailed description of these criteria), chosen to enroll into the study, and been randomized into the treatment (MARSSI) or control condition (podcast group*)*. The act of randomization constitutes the offer to participate and is the point at which the individual becomes a participant in the study. The analytic sample for our primary research questions will be all participants who were randomized into either the treatment or control conditions and who have reported the necessary outcome data to construct the primary outcome measure of interest. We will impute missing baseline data for any participants in this group who are missing either baseline covariate data and/or baseline outcome data. Missing data procedures are outlined in *Data cleaning, subsection iv* above.

d. Assessment of baseline equivalence

Baseline equivalence will be reported for all baseline measures of each outcome variable as well as relevant demographic and sexual behavioral measures. We first list and describe the measures we will use to examine the equivalence of our treatment and control groups at baseline. After we identify the measures, we provide details on the diagnostic methods that we will use to assess any baseline differences that may exist between the treatment and control groups in the measures outlined below.

Demographic Measures

Baseline equivalence will be assessed for four demographic variables. Age is constructed from data gathered by PRG study coordinators in the *Eligibility Screening Form*. Race, gender, and ethnicity are constructed using participant self-responses to questions in the baseline *Participant Questionnaire*. For the race variables, categorical responses to a single question are used to create multiple dichotomous variables. We provide details on variable coding below; details on variable construction can be found in Table 4.

Demographic:

- Age at screening (continuous; range 16-21)[28]
- Gender[29]
- Race[30]
- Hispanic/Latino/a

[28] Age at screening is calculated using the individual's reported date of birth.

[29] At baseline, participants are asked "What is your gender?" and provided with a list of the following response options: *Female; Male; Transgender female; Transgender male; Non-binary/genderqueer; Unsure/questioning; I do not identify as any of these (Would you like to indicate how you identify yourself?)*. A dummy variable is generated using the data reported for this item, where 1= participants who self-identified as female at baseline and 0= participants who self-identified as one of the other categories at baseline.

[30] At baseline, participants are asked "What is your race/ethnicity?" and are provided with a list of the following response options: *American Indian or Alaskan Native; Asian; Black or African American; Hispanic or Latino/a; Native Hawaiian or other Pacific Islander; White;* or *Some other race/ethnicity (please specify)*. Participants can select more than one category and they can also specify some other race/ethnicity. This item is used to create one dummy variable (White). For this dummy variable, individuals are coded as 1 if they self-identified only as "White" and 0 if otherwise.

Baseline Outcome Measures

In addition to the demographic variables, we will assess baseline equivalency of baseline measures of the outcome measures. We provide details on variable coding below; details on variable construction can be found in Table 2.

- *Frequency of having vaginal sex without condoms in the past 30 days* at baseline (continuous; values range 0 to *k*, where 0= has had vaginal sex without condoms 0 times in past 30 days and *k*= number of times having vaginal sex without condoms in past 30 days)
- *Current use of effective non-barrier contraception* at baseline (dichotomous; values of 0 or 1, where 0= not currently using effective non-barrier contraception and 1= currently using effective non-barrier contraception)
- *Number of sexual partners in the past 3 months* at baseline (continuous; values range 0 to *k*, where 0= has 0 sexual partners in past 3 months and *k*= number of sexual partners in past 3 months)

Balance Assessment Methods

We propose to assess baseline equivalence of the treatment and control groups according to a multi-step procedure. Baseline equivalence statistics will be produced for each analytic sample.[31] Only participants who provide sufficient baseline and primary outcome data (i.e., non-missing) will be included in the test for baseline equivalence of the analytic sample for the specified outcome; thus, the analytic sample used for each research question may vary slightly because of the exclusion of non-responders.[32] As required by the "Identifying Programs that Impact Teen Pregnancy, Sexually Transmitted Infections, and Associated Sexual Risk Behaviors" review protocol version 6.0, we will report the standardized mean difference of each baseline variable for the treatment and control groups.[33]

To establish baseline equivalence, we propose to generate model-based point estimates of the difference between the treatment and control group means for the identified baseline equivalence variables. We will report the adjusted means of the difference in adjusted means of the baseline variable of interest for the treatment and control groups. We will then compute the pooled standard deviation of these variables. Finally, we will produce a standardized difference of means by dividing the first term by the second. The steps for establishing baseline equivalence using standardized mean difference are outlined below:

[31] Due to item-missing outcome data, we expect there may be slight differences in analytic samples for each research question.

[32] Note that our benchmark approach is to produce diagnostic estimates of baseline equivalence on the sample of observations that have non-missing baseline and outcome data, without any adjustments to missing baseline data. We will conduct a sensitivity test, however, that calculates baseline equivalence using the exact same samples of observations that we will use in our primary analysis by applying the missing data approach outlined in *Data cleaning, subsection iv*.

[33] Mathematica Policy Research. (2022). Identifying Programs that Impact Teen Pregnancy, Sexually Transmitted Infections, and Associated Sexual Risk Behaviors: Review Protocol, Version 6.0.

step 1. First, we generate a model-based estimate of the difference between treatment and comparison groups on the baseline measures identified above. Separate models will be run for each of the baseline variables. The empirical model will be estimated with OLS (using Stata). If the measure is dichotomous, we propose to use a linear probability model to estimate the predicted probability of group membership. The model is a reduced-form variation of the model that we use to estimate program impact (as detailed in the *Model specification and covariates* section, below).[34]

$$Y_{baseline} = \beta_0 + \beta_1 T + \sum(\beta_n D_n) + \varepsilon$$

where:

$Y_{baseline}$ – is the baseline measure of the variable that we use to establish baseline equivalency. This variable is included as a covariate in the benchmark analytic model (see Table 4 for details on variable coding). Separate models will be estimated for each baseline measure specified above.

β_0 – The intercept term, which represents the adjusted mean value of the baseline equivalency measure for participants in the control sample, with all other variables in the model held constant at zero.

β_1 – This represents the adjusted (but not standardized) mean difference in the baseline equivalency variable between treatment and control participants.

T - A dummy treatment indicator variable whose value equals 1 if the participant is randomized into the treatment group and zero otherwise.

D – An n-1 vector of blocking variables (i.e., subgroups within which random assignment occurred), where n represents the full enumeration of all state by coordinator blocks from which participants were randomly assigned to a condition. For each of these n coordinator-by-state blocks, we include in the baseline equivalence model a dummy indicator variable that will identify whether a study participant was randomized within that block (1) or not (0).

ε – The residual or random variation that remains for each observation after the structural components of the model are estimated. It is the difference between the observed and the predicted values at the individual level.

step 2. Report the adjusted means of the differences in the baseline variable of interest for the treatment and control groups.

step 3 (continuous variables only). If the baseline measure is continuous, we propose to use the following formula to calculate the pooled within-group standard deviation of the outcome measure:

[34] It is a reduced-form because individual-level, demographic covariates are omitted. It is a variation because the dependent variable is the baseline equivalence variable, not the outcome measure.

$$S_p = \sqrt{\frac{(n_t - 1)S_t^2 + (n_c - 1)S_c^2}{(n_t + n_c - 2)}}$$

where: n_t and n_c are the sample sizes, and S_t and S_c are the participant-level standard deviations for the baseline measures for the analytic treatment and comparison groups, respectively. We will produce separate calculations of the pooled standardized deviation for each variable used to establish baseline equivalence (as noted above).

step 4. Produce the standardized difference of means. If the pre-intervention measure is continuous, we will use the formula for Hedges' *g* to compute the standardized difference of means for the treatment and comparison groups:

$$g = \frac{\beta_1}{S_p}$$

where: β_1 is the adjusted mean difference in the variable selected to establish baseline equivalence for the treatment and comparison groups (calculated in Step 1), and S_p is the pooled standard deviation (produced in Step 3).

For dichotomous baseline variables we will use the Cox index, which yields effect size values similar to the values of Hedges' *g* that one would obtain if group means, standard deviations, and sample sizes were available, assuming the dichotomous outcome measure is based on any underlying normal distribution. Following this guidance, we propose to use the Cox index to estimate baseline equivalence for dichotomous baseline covariates. The formula is as follows:

$$d_{Cox} = \left[ln\left(\frac{p_t}{1 - p_t}\right) - ln\left(\frac{p_c}{1 - p_c}\right) \right] \Big/ 1.65$$

where: pt and pc represent the probability of occurrence of the event (or characteristic) within the treatment and comparison groups, respectively.

e. Benchmark analytic approach for primary research questions

As detailed in our primary research questions, this study investigates whether offering MARSSI to participants impacts their reported frequency of having vaginal sex without condoms, use of effective non-barrier contraception, and number of sexual partners. We do this within the intent to treat (ITT) framework, which does not measure the effect of the participant's exposure to the treatment itself but rather the effect of the offer of the treatment relative to the offer of receiving the control condition. This framework maintains the integrity of the experimental structure by including all participants who were randomized (except those who attrite) in the analytic sample, thereby maintaining an exogenous assignment of participants to experimental condition. Bias can be insinuated, however, through self-selection if any participant who is randomized fails to provide outcome data.

i. Estimation approach

The primary research questions under investigation in this study are whether offering MARSSI to participants impacts their: 1) reported frequency of having vaginal sex without condoms, 2) use of effective non-barrier contraception, and 3) reported number of sexual partners (see Table 2 for variable constructions).

We propose to estimate these impacts using a regression-based approach that will model intervention effects as a function of assignment to MARSSI (i.e., treatment), relevant baseline covariates, a baseline measure of the outcome variable, and the combined state and coordinator-level (dummy blocking) indicators (see Table 4 for variable constructions).[35] In addition, missing baseline data indicators will be included in the model for each baseline variable in which missing values are imputed through dummy variable adjustment. Although a straight difference-of-means approach should provide unbiased estimates of the effect of the treatment, we propose a model-based approach because it will increase the precision of those estimates and purge any small differences associated with baseline imbalance. The empirical model will be estimated with an OLS regression (using Stata). We present the empirical model here:

$$Y_{Post} = \beta_0 + \beta_1 T + \beta_2 Y_{Pre} + \sum(\beta_P X_P) + \sum(\beta_n D_n) + \sum(\beta_p M_p) + \varepsilon$$

where:

Y_{Post} - The outcome of interest, either: 1) times having vaginal sex without condoms in the past 30 days (continuous; values range 0 to *k*, where 0= has had sex without condoms 0 times in past 30 days, and *k*= number of times having vaginal sex without condoms in past 30 days); 2) number of sexual partners in the past 3 months (continuous; values range 0 to *k*, where 0= has 0 sexual partners in past 3 months, and *k*= number of sexual partners in past 3 months); or 3) current use of effective non-barrier contraception (dichotomous; where 1= currently using effective non-barrier contraception and 0= not currently using effective non-barrier contraception) reported by participant *i* three months post-intervention (see Table 2 for full details on the variable construction).

β_0 – The intercept term, which represents, depending on the outcome measure of interest in the analysis, the outcome for the average control participant with all other variables in the model held constant at zero.

β_1 – This is the parameter estimate of substantive interest. β_1 represents, depending on the outcome measure of interest in the analysis, either: 1) the

[35] With the assumption that we maintain low overall and differential attrition and that the study otherwise executes the RCT with integrity, we should be able to estimate an un-biased estimate of the average treatment effect of MARSSI by comparing differences in the means of the outcome variable reported by the treatment group with those reported by the control group. We could then provide a compelling response to our research question by testing the hypothesis that there is no difference between the two groups using straight-forward hypothesis testing statistics (t-test). With that said, we propose a regression-based model that includes covariates, because randomization should ensure covariates are uncorrelated with the treatment variable (i.e., they should not affect the estimate of the treatment effect), and the inclusion of covariates may improve the precision/efficiency of the estimate of the treatment effect. See: Angrist, J. D., & Pischke, J. (2009). Mostly harmless econometrics: An empiricist's companion. Princeton: Princeton University Press; Rosenblum, M. and van der Laan, M. J. (2009), Using Regression Models to Analyze Randomized Trials: Asymptotically Valid Hypothesis Tests Despite Incorrectly Specified Models. Biometrics, 65: 937-945. doi:10.1111/j.1541-0420.2008.01177.x. Robinson, L.D. & Jewell, N.P. (1991). Some Surprising Results About Covariate Adjustment in Logistic Regression Models. International Statistical Review, 58(2), 227-240.

adjusted mean difference between treatment and control participants' self-reported times having vaginal sex without condoms in the past 30 days three months post-intervention; 2) the adjusted mean difference between treatment and control participants' self-reported number of sexual partners in the past 3 months three months post-intervention; or 3) the adjusted mean difference between the proportion of treatment participants who self-report using effective non-barrier contraception and control participants who self-report using effective non-barrier contraception three months post-intervention.

T –A dummy treatment indicator variable whose value equals 1 if the participant is randomized into the treatment group and 0 otherwise.

Y_{Pre}– The baseline measure of the outcome variable of interest reported by participant *i* at baseline (see Table 2 for full details on the variable construction); variable will be re-centered at the grand mean for analysis.

X – A p vector of baseline (i.e., measured prior to receiving intervention or exogenous to treatment) participant-level covariates to account for the variation in outcomes associated with these groups. These covariates, listed in detail in Table 4, will include:

a) Age – self-reported age (based on date of birth) at screening (continuous; range 16-21); variable will be re-centered at the grand mean for analysis.
b) Gender – gender of participants as self-reported at baseline (coded as 1 if female and coded as 0 if otherwise); variable will be re-centered at the grand mean for analysis.
c) Race – race of participant as self-reported at baseline (coded as 1 if participant self-identified as only White and 0 if otherwise); variable will be re-centered at the grand mean for analysis.
d) Ethnicity – Hispanic/Latino(a) ethnicity of participant as self-reported at baseline (coded as 1 if identify as Hispanic/Latino(a) and 0 if do not identify as Hispanic/Latino(a)); variable will be re-centered at the grand mean for analysis.

D – An n-1 vector of blocking variables (i.e., subgroups within which random assignment occurred), where n represents the full enumeration of all state-by-coordinator blocks from which participants were randomly assigned to condition. For each of these n coordinator-by-state blocks, we include in the analytic model a dummy indicator variable that will identify whether a study participant was randomized within that block (1) or not (0). The variables will be re-centered at the grand mean for analysis.

M – A p vector of missing baseline data indicator variables representing each of the p baseline covariates that had missing observations (coded as 1 if the observation for that variable is missing and 0 if it is non-missing).

Using a vector of dummy variables (LSDV) in a regression to account for a blocked randomization scheme is a conventional choice for a situation where each block has the same probability of assignment.[36] This is consistent with our design specification and so we have adopted the dummy variable modeling strategy here. As Angrist (1998) and others have pointed out, however, scenarios exist where the strategy may fail to identify the estimand of interest for a variety of reasons. Of particular interest

[36] Gerber, A. S., & Green, D. P. (2012). Field Experiments: Design, Analysis and Interpretation. W.W. Norton & & Company.

here is that the dummy-blocking scheme may exclude blocks/cases from the full randomized sample. Regardless of what the nominal randomization probabilities may be, one feature of the blocking model is that it will assign zero weight to any block partition where the actualized probability of treatment equals 0 or 1. That is, study assignment blocks that contain only a single participant (who is then assigned to either treatment or control) are functionally excluded from the treatment effect estimate.

If this is pervasive enough in the analytic sample – where multiple cells exist with just a single participant – impact estimates may fail to identify the treatment effect for the fullest possible ITT sample. In this case we will alter our benchmark approach and substitute the LSDV procedure with an almost-equivalent inverse-probability-of-treatment-weighted (IPTW) regression, which should mitigate the exclusion problem, while still weighting cells accordingly. This estimation approach is almost equivalent to the LSDV blocking procedure except that it is functionally saturated with respect to the treatment and dummy variables (but not covariates), and it can be constructed to broaden the analytic sample to include cells where only a single case was randomized and (calculated) probability of treatment is 0 or 1.[37] When we use a normalized IPTW and estimate the propensity score with the linear probability model he single-case blocks are given reduced but not zero weight.

If we use the IPTW regression as the benchmark approach, we will weight each observation by the inverse probability of assignment for their condition in their block. For the treatment group, the weights will be $\frac{1}{e(x)}$; for the control group, weights will be $\frac{1}{1-e(x)}$. Where e(x) is the propensity score estimated by the by the linear probability model using the following specification:

$$e(x) = \alpha_0 + \sum(\beta_n D_n) + \varepsilon$$

Prior to estimation we will normalize weights such that the sum of all weights will be equal to 1.

We will report model-estimated effects and the results of significance tests in the findings section of the final impact report. Statistical significance will be based on test statistics produced by Stata for the coefficient β_1 using a two-tailed test, with $p < .05$.

ii. Adjustment for baseline differences

As described in the *Estimation approach* subsection above, we will include covariates for the identified demographic characteristics and a baseline measure of the outcome of interest in each our benchmark models to increase the precision of our estimates and account for any small baseline differences between the treatment

and control groups.

iii. Additional covariates

We do not intend to include any further covariates beyond those specified in the *Estimation approach* subsection above.

Table 4. Covariates included in impact analyses

Covariate	Description of the covariate
Baseline primary outcome measures	
Frequency of vaginal sex without condoms in the past 30 days	The risk outcome is operationalized as the number of times in the past 30 days a person reports having vaginal sex without using a condom. The measure is calculated from the following item on the *Participant Questionnaire*: • In the past 30 days, how many times have you had vaginal sex without using a condom? The resulting variable is continuous with values that range from 0 to *k*, where 0 indicates that a person has not engaged in vaginal sex without a condom in the past 30 days, and *k* indicates the number of times the person has engaged in vaginal sex without a condom (risk behavior) in the past 30 days. Note: All respondents who have baseline data and have provided a response to either this question or have indicated they have not had vaginal sex in the past 30 days will be included in the construction of this measure. Persons who indicate that they have not had vaginal sex in the past 30 days are coded as having vaginal sex without a condom zero times.
Current use of effective non-barrier contraception	The protective outcome is operationalized as a dichotomous variable indicating whether a person reports currently using effective non-barrier contraception or not. The measure is calculated from the following item on the *Participant Questionnaire*: • Please indicate which method of prescription birth control you currently using: o Oral contraceptives (for example, the pill) o The patch (for example, Ortho Evra) o The shot/injection (for example, Depo-Provera) o The ring (for example, NuvaRing) o The implant (for example, Implanon or Nexplanon) o IUD (for example, ParaGard, Skyla, Mirena, Kyleena, Liletta) o None of the above A person who selects *Oral contraceptives, patch, shot/injection, ring, implant,* or *IUD* is given a value of 1 for the measure. A person who selects *None* is given a value of 0 for the measure. The resulting variable is dichotomous with values 0 or 1, where 0 indicates a person who does not currently use effective non-barrier contraception and 1 indicates a person who does currently use effective non-barrier contraception. Note: All respondents who have baseline data and have provided a response to this question will be included in the construction of this measure.
Number of sexual partners in the past 3 months	The risk outcome is operationalized as the number of sexual partners in the past three months. The measure is calculated from the following item on the *Participant Questionnaire*: • How many sexual partners have you had in the past 3 months? The resulting variable is continuous with values that range from 0 to *k*, where 0 indicates that a person has had no sexual partners in the past three months, and *k* indicates the number of sexual partners in the past three months. Note: All respondents who have baseline data and have provided a response to this question will be included in the construction of this measure.

Individual-level covariates	
Age	The variable is measured as the respondent's self-reported age in years at screening. The measure is constructed from the following item on the *Eligibility Screening Form*: • Date of birth The variable is calculated by subtracting the reported date of birth given from the date when the screening was completed. The resulting variable is continuous with values ranging from 16 to 21.
Gender	The measure is operationalized a dummy variable, where 0 = identify otherwise*;* 1 = identify as female. The measure is taken from the following item on the baseline *Participant Questionnaire:* • What is your gender? o Female o Male o Transgender female o Transgender male o Non-binary/genderqueer o Unsure/questioning o I do not identify as any of these (Would you like to indicate how you identify yourself? _____) Variable will be coded as 1 if participant self-identified as female; if female is not selected, the response will be coded as 0.
Race	The measure is operationalized as a dummy variable, where 0 = identify as another race and/or ethnicity; 1 = identify only as white. The measure is taken from the following item on the baseline *Participant Questionnaire:* • What is your race/ethnicity? (Participants can select more than one response) o American Indian or Alaska Native o Asian o Black or African American o Hispanic or Latino/a o Native Hawaiian or other Pacific Islander o White o Some other race (specify)_____ Variable will be coded as 1 if participant self-identified as white only; if another race and/or ethnicity is selected or the respondent identifies as multiracial, the response will be coded as 0.
Hispanic /Latino/a	The measure is operationalized as a dummy variable, where 0 = identify as another ethnicity/do not identify ethnicity; 1 = identify as Hispanic or Latino/a. The measure is taken from the following item on the baseline *Participant Questionnaire:* • What is your race/ethnicity? (Participants can select more than one response) o American Indian or Alaska Native o Asian o Black or African American o Hispanic or Latino/a o Native Hawaiian or other Pacific Islander o White

	o Some other race (specify) Variable will be coded as 1 if participant self-identified as *Hispanic or Latino/a*, regardless as to whether other races/ethnicities are specified; if Hispanic/Latino/a origin is not selected, the response will be coded as 0.
Blocking covariates	
Study coordinato by State blocks	A set of n-1 dummy variables, where *n* refers to the full set of study coordinator by state blocks within which participants are randomly assigned to the treatment or control condition. Participants are enrolled within one of 255 possible blocks defined by (1) the research coordinator who enrolled the participant into the study and (2) the state (or District of Columbia) where the participant indicated they would be most likely to seek reproductive health care. To clarify, each of the 5 research coordinators assign participants within one of their own 51 separate (state-and DC-based) random allocation lists, for a total potential 255 random allocation blocks. Data for the measure are obtained from the *Enrollment Log database* (coordinator) and the *Eligibility Screening Form* (state). Each dummy will be coded as 1 if the individual was jointly enrolled by a particular study coordinator and indicated they would be most likely to seek reproductive health care in a particular state, and 0 if otherwise. Dummy variables will be grand mean-centered so that the intercept will then reflect the un-weighted mean study coordinator*state effect.

f. Analytic approach for secondary research questions

For the secondary research questions listed under the subsection *Analysis of secondary outcomes in full sample using ITT approach* in the *Impact Study Research Questions* section, we will use the same analytic approach as described above under the *Benchmark analytic approach for primary research questions* subsection. For the secondary research questions listed under the subsection *Analysis of primary and secondary outcomes in subgroups using ITT approach*, models will also include a measure for the subgroup of interest and an interaction term for the subgroup and treatment to assess differences in subgroup impacts. For the secondary research questions listed under the subsection *CACE analysis of primary and secondary outcomes*, we will use CACE analysis. For this, we will assess the effect of MARSSI when individuals participate in (comply with) the program using two analytic methods derived from the principal stratification framework that estimate impact conditioned on endogenous (post-randomization) compliance: principal score weighting and two-stage least squares regression.[38] We will use the CACE methods described below when assessing these questions in both the full study sample and specified subgroups.

i. Principal score method

The first method we will use to estimate the CACE is a balancing procedure, based on propensity score methods, which can be used in settings where principal stratum membership (compliance) is observable under one treatment condition. The use of propensity scores is a common technique to balance treatment and comparison groups in nonexperimental studies; however,

[38] Frangakis, C. E., & Rubin, D. B. (2002). Principal stratification in causal inference. Biometrics, 58(1), 21–29. https://doi.org/10.1111/j.0006-341x.2002.00021.x

propensity score methods have also been used to address noncompliance in RCTs as well.[39]

Within the context of a randomized trial using a 1:1 assignment ratio, there is an expectation that principal stratum membership should be equally allocated to the treatment and control groups. In other words, if there are compliers in the treatment group, we would expect there exists a similar group of individuals in the control group who would have complied if the program had been offered to them. Whereas the conventional use of propensity scores aims to model treatment group membership (where treatment group membership is the same as intervention receipt), the aim here is to use propensity scores to model treatment receipt (compliance) in the treatment group and subsequently predict probability of principal stratum membership among members of the control group.

In accordance with Hill et al. (2002), to distinguish this latter prediction step for the control group, we refer to their probability of principal stratum membership as the principal score.[40] The core assumption in propensity score methods is that of conditional ignorable treatment assignment, or the assertion that treatment assignment is independent of the potential outcomes, given a set of observed covariates (Rosenbaum & Rubin, 1983). When we use a probability score to balance treatment and control groups to estimate CACE in an RCT, this assumption now applies to principal stratum membership (compliance). In other words, principal stratum membership is independent of the potential outcomes given the observed set of covariates (Jo & Stuart, 2009).

We will follow the steps outlined in Stuart & Jo (2015) to estimate the CACE using principal score weights. Briefly, these will include: 1) using the same baseline covariates used in the benchmark analysis to predict compliance among the treatment group; 2) predicting the probability of compliance (principal score) among members of the control group; 3) creating analytic weights reflecting the probability of compliance; and 4) estimating CACE by fitting the outcome model using the principal score weights. Consistent with the ITT analysis, we will use an appropriate regression model (OLS for continuous/scale measures, logit for dichotomous, count for frequency) for each outcome. We will include the following covariates in each model: age, gender, race, ethnicity, randomization blocks, and the baseline measure of the outcome.

[39] Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika,70(1), 41–55.https://doi.org/10.1093/biomet/70.1.41. Follmann, D. A. (2000). On the effect of treatment among would-be treatment compliers: An analysis of the multiple risk factor intervention trial. Journal of the American Statistical Association,95(452), 1101–1109.https://doi.org/10.1080/01621459.2000.10474306. Jo, B., & Stuart, E. A. (2009). On the use of propensity scores in principal causal effect estimation. Statistics in Medicine,28(23), 2857–2875.https://doi.org/10.1002/sim.3669. Stuart, E. A., & Jo, B. (2015). Assessing the sensitivity of methods for estimating principal causal effects. Statistical Methods in Medical Research,24(6), 657–674.https://doi.org/10.1177/0962280211421840.

[40] Hill, J., Waldfogel, J., & Brooks-Gunn, J. (2002). Differential effects of high-quality child care. Journal of Policy Analysis and Management,21(4), 601–627.https://doi.org/10.1002/pam.10077.

ii. Instrumental variable method

The second method we will use is an instrumental variable approach that uses the random assignment mechanism to act as an instrument for compliance to estimate the CACE. We will produce the CACE with a joint model that first estimates participation, given treatment assignment and subsequently estimates the outcome, given participation; this is known as Two-Stage Least Squares (TSLS) regression.[41] Instrumental variable analysis is a common technique in evaluation to estimate the CACE in randomized trials.[42]

We will estimate the CACE with the *ivregress 2sls* command in Stata 16.1 (StataCorp, 2019).The first stage model predicts compliance (full participation) using the instrument (treatment assignment). The second stage predicts the outcome, given participation.

This simultaneous estimation framework allows the user to calculate accurate standard errors that account for the uncertainty in the first stage model (Stuart et al.,2008).

The benefit of the TSLS model is that it allows for the inclusion of baseline covariates that predict both participation and the outcome, which can help further reduce the amount of error in the estimation and possibly reduce bias due to exclusion restriction violations.[43] As with the benchmark approach and principal score method approach, we will include the following covariates in both stages: age, gender, race, ethnicity, a baseline measure of the outcome of interest, and a series of dummy variables representing the randomization blocks.

g. Sensitivity analyses

We will conduct sensitivity analyses to test the robustness and validity of our benchmark approaches outlined above. These include: 1) estimating program effects using OLS regression without baseline covariates and blocking variables; 2) implementing alternative models to test for bias; 3) excluding baseline covariates; 4) not imputing or adjusting for missing data; 5) excluding unreliable data; 6) excluding outliers; 7) condensing data collection windows to exclude late responders; and 8) using alternative model specifications to estimate program effects.

[41] Angrist, J., & Imbens, G. W. (1995). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. Journal of the American Statistical Association,90(430), 431–442.https://doi.org/10.2307/2291054.
[42] Connell, A. M. (2009). Employing complier average causal effect analytic methods to examine effects of randomized encouragement trials. The American Journal of Drug and Alcohol Abuse,35(4), 253–259.https://doi.org/10.1080/00952990903005882. Schochet, P. Z., & Chiang, H. S. (2011). Estimation and identification of the complier average causal effect parameter in education RCTs. Journal of Educational and Behavioral Statistics,36(3), 307–345.https://doi.org/10.3102/1076998610375837. Stuart, E. A., Perry, D. F., Huynh-Nhu, L., & Ialongo, N. S. (2008). Estimating intervention effects of prevention programs: Accounting for noncompliance. Prevention Science,9(4), 288–298.https://doi.org/10.1007/s11121-008-0104-y.
[43] Jo, B. (2002). Statistical power in randomized intervention studies with noncompliance. Psychological Methods,7(2), 178–193.https://doi.org/10.1037/1082-989x.7.2.178.

i. **Difference of means.** Our benchmark approach is to include baseline covariates and blocking variables in an OLS regression model to improve the precision of our effect estimates. Assuming our sample is sufficiently large, this approach should generate unbiased estimates.[44] To test this, we will conduct sensitivity analyses that use OLS and only the treatment indicator included as an independent variable. This approach should approximate a difference of means t-test. We will report any substantive differences in the results section of the final manuscript and analytic findings for both approaches will be presented alongside each other in an appendix.

ii. **Alternative models.** Our benchmark approach includes covariates and accounts for blocking procedures either using fixed effects for the study coordinator-by-state assignment blocks or an IPTW approach that is similar but does not zero-weight single-case cells. This strategy offers a compromise between the unbiasedness of the difference-of-means approach and the added precision and statistical power offered by regression adjustment. Heterogeneous treatment effects, however, remain a possible threat to this approach. As such, we propose to fit a model that constructs a treatment effect from a weighted average of a fully saturated OLS regression model. We will use the weighted least squares version of this model proposed by Lin (2013, p.10). We will report any substantive differences in the results section of the final manuscript and analytic findings for both approaches will be presented alongside each other in an appendix.

iii. **Without baseline covariates.** Our benchmark approach is to include baseline covariates in our model to improve the precision of our estimates. To test this, we will conduct sensitivity analyses that involve running identical empirical models without the baseline covariates included. We will report any substantive differences in the results section of the final manuscript and analytic findings for both approaches will be presented alongside each other in an appendix.

iv. **Without adjusted baseline data.** As outlined in the *Missing data approach* section, our benchmark approach is to adjust baseline data as published guidance suggests that this may produce unbiased impact estimates and maximize the use of available data. We will test this by way of sensitivity analyses that involve running identical empirical models without the adjusted data. As outlined in the *Assessment of baseline equivalence* section, we will also produce diagnostic estimates of baseline equivalency on the baseline outcome variables according to our benchmark approach. We will report any substantive differences identified in these analyses in the results section of the final manuscript and analytic findings for both approaches will be presented alongside each other in an appendix.

v. **Without unreliable data.** As discussed in the *Data cleaning* section, data for cases that are deemed unreliable are flagged, but still included in benchmark analyses. We will also conduct sensitivity analyses that involve running identical empirical models with the unreliable data excluded. We will report any substantive differences in the

[44] Angrist, J. D., & Pischke, J.-S. (2009). Mostly harmless econometrics: An empiricist's companion. Princeton: Princeton University Press; Lin, W. (2013). "Agnostic Notes on Regression Adjustments to Experimental Data: Re-Examining Freedman's Critique" The Annals of Applied Statistics. 7(1) 295-318.

results section of the final manuscript and analytic findings for both approaches will be presented alongside each other in an appendix.

vi. **Without outliers**. As discussed in the *Data cleaning* section, extreme data values are investigated and flagged as outliers. Our benchmark analytic approach is to include data flagged as outliers (i.e., extreme values that are not considered invalid) in analysis. We will also conduct sensitivity analyses that exclude these data. We will report any substantive differences in the results section of the final manuscript and analytic findings for both approaches will be presented alongside each other in an appendix.

vii. **Condensed data collection windows**. Our benchmark approach is to include follow-up data from all participants who completed a questionnaire during their open data collection window, regardless of the time point in that window when it was completed. Data collection windows are broad to minimize attrition from the analytic sample. To examine whether or not this influences our results – and, in particular, whether or not study participants who respond later report different outcomes from those who respond earlier – we will conduct an analysis that examines the difference, if any, in response time between treatment and control participants and compares impact estimates for analytic samples without late responders. Late responders will be defined as those participants who complete their long-term questionnaire more than one month after the initiation of the three-month data collection window. We will report any substantive differences in the results section of the final manuscript and analytic findings for both approaches will be presented alongside each other in an appendix.

viii. **Statistical modeling.** We have proposed using OLS regression as the benchmark statistical method for producing impact estimates. We will conduct analyses to test robustness of this choice and to assess whether there are substantive differences in the point estimates of interest produced by OLS and alternative estimators.[45] Specifically, for each research question, we will compare OLS estimates with those derived from models that may fit the distribution of the data better. For the two count outcomes (frequency of having vaginal sex without condoms and number of sexual partners), the treatment effect will be estimated with an appropriate count model (using Stata); for the dichotomous outcome, the effect will be estimated using a logit model.[46] If there is a substantive difference in the impact estimates of interest, we will report the results of each.

h. Bayesian interpretation

In addition to assessing our primary research question findings using a traditional frequentist approach, we will supplement our presentation of impact estimates with a Bayesian interpretation based on posterior probabilities of program effectiveness. To do

[45] To account for potential violations of model assumptions (e.g., heteroskedasticity, overdispersion), we will use robust standard errors in all analyses.

[46] For count models, we will assess model fit using diagnostic model-fit methods, including the Stata command countfit, AIC and BIC model fit statistics, and by assessing predicted versus observed count probabilities for competing models. See Hilbe, J. M. (2014). Modeling count data. New York: Cambridge University Press.

this, we will rely upon the BAyesSian Interpretation of Estimates (BASIE) framework.[47] The prior distributions for our Bayesian estimate will be informed by the meta-analysis of teen pregnancy prevention intervention program effects conducted by Juras et al.[48] We will calculate the precision-weighted average of the traditional estimate (i.e., shrunken estimate) based on this prior evidence using the BASIE probability tool provided by Deke et al. The shrunken estimates will be treated as sensitivity analyses to our traditional estimates. We will report the shrunken estimates alongside the estimates derived from our impact study data and interpret them using posterior probabilities.

[47] Deke, J., Finucane, M., & Thal, D. (2022). The BASIE (BAyeSian Interpretation of Estimates) Framework for Interpreting Findings from Impact Evaluations: A Practical Guide for Education Researchers. Toolkit. NCEE 2022-005. *National Center for Education Evaluation and Regional Assistance*.

[48] Juras, R., Kelsey, M., Steinka-Fry, K., Lipsey, M., Layzer, J., & Tanner-Smith, E. (2022). Meta-analysis of Federally Funded Adolescent Pregnancy Prevention Program Evaluations. *Prevention Science*, *23*(7), 1169-1195.

5. Additional planned analyses[49]

In addition to the primary and secondary research questions described above, we intend to investigate the effect of mediating factors on primary outcomes of interest and one secondary outcome of interest.

Effects of Mediators on Primary and Secondary Outcomes of Interest

a. What are the are the short-term (three months post-intervention) and long-term (six months post-intervention) impacts of the offer to participate in MARSSI (treatment) relative to the offer to participate in the control condition on participants' reported *frequency of having vaginal sex without condoms* considering the following potential mediators:
    a. Sexual communication self-efficacy
    b. Condom planning self-efficacy
    c. Condom knowledge
    d. Contraceptive knowledge
b. What are the are the long-term (six months post-intervention) impacts of the offer to participate in MARSSI (treatment) relative to the offer to participate in the control condition on participants' reported *frequency of having vaginal sex without condoms* considering the following potential mediators:
    a. Motivation to use condoms
c. What are the are the short-term (three months post-intervention) and long-term (six months post-intervention) impacts of the offer to participate in MARSSI (treatment) relative to the offer to participate in the control condition on participants' reported *use of effective non-barrier contraception* considering the following potential mediators:
    a. Sexual communication self-efficacy
    b. Contraceptive planning self-efficacy
    c. Contraceptive knowledge
d. What are the are the long-term (six months post-intervention) impacts of the offer to participate in MARSSI (treatment) relative to the offer to participate in the control condition on participants' *use of effective non-barrier contraception* considering the following potential mediators:
    a. Motivation to use prescription birth control
e. What are the are the short-term (three months post-intervention) and long-term (six months post-intervention) impacts of the offer to participate in MARSSI (treatment) relative to the offer to participate in the control condition on participants' reported *number of sexual partners* considering the following potential mediators:
    a. Sexual communication self-efficacy
    b. Condom planning self-efficacy
    c. Contraceptive planning self-efficacy
    d. Coping self-efficacy
    e. Condom knowledge
    f. Contraceptive knowledge

[49] Note that we have not included a research question examining how COVID-19 may have influenced program implementation and participant outcomes because the MARSSI intervention was implemented exclusively in a virtual format during the full course of the study. As such, we are not able to explore differences between in-person and online implementation.

f. What are the are the long-term (six months post-intervention) impacts of the offer to participate in MARSSI (treatment) relative to the offer to participate in the control condition on participants' reported *number of sexual partners* considering the following potential mediators:
    a. Motivation to purposefully abstain from sex
g. What are the are the short-term (three months post-intervention) and long-term (six months post-intervention) impacts of the offer to participate in MARSSI (treatment) relative to the offer to participate in the control condition on participants' reported *depressive symptoms* considering the following potential mediators:
    a. Coping self-efficacy

These questions pertaining to potential mediators will be explored only if the main relationship between treatment and the primary outcome of interest is found to be statistically significant at the $p < .05$ level. In this scenario, we will conduct mediation analysis to estimate the total, indirect, and direct effects of treatment on the outcomes listed above, considering the specified mediators, using an appropriate modeling approach based on the type of outcome being explored.

Appendix A: Participant Questionnaire

BEAM Health Study
Participant Questionnaire

[authenticator] Please enter your ID number: _____

Please re-enter your ID number: _____

SECTION 1: BACKGROUND INFORMATION

In this first section, we'd like to get some general information about you, your family, and your education. This information is used only for reporting purposes to describe the types of individuals completing this questionnaire. Your name is NOT on this questionnaire, and the information you provide will be kept completely confidential. Please be honest in your responses.

q1 **What is your race/ethnicity? Select all that apply.**

- ☐ American Indian or Alaska Native
- ☐ Asian
- ☐ Black or African American
- ☐ Hispanic or Latino/a
- ☐ Native Hawaiian or other Pacific Islander
- ☐ White
- ☐ Some other race/ethnicity (please specify): _____

q2 **What is your gender? Select one answer.**

- ☐ Female
- ☐ Male
- ☐ Transgender female
- ☐ Transgender male
- ☐ Non-binary/genderqueer
- ☐ Unsure/questioning
- ☐ I do not identify as any of these (Would you like to indicate how you identify yourself? _____)

q3 **Are you (or your family) eligible for Medicaid or a similar state-sponsored health insurance plan?**

- ☐ Yes
- ☐ No
- ☐ Don't know

q4 **What grade or level of school are you currently in? If you are enrolled in school but between grades right now, select the most recent grade or level of school that you completed.**

- ☐ Grade 1
- ☐ Grade 2
- ☐ Grade 3
- ☐ Grade 4
- ☐ Grade 5
- ☐ Grade 6
- ☐ Grade 7
- ☐ Grade 8
- ☐ Grade 9
- ☐ Grade 10
- ☐ Grade 11
- ☐ Grade 12
- ☐ GED program
- ☐ Technical/vocational training/college
- ☐ Ungraded (school without formal grade levels)
- ☐ Not currently in school

q5 **What is the highest grade in school that your mother finished?**

- ☐ Did not finish high school
- ☐ Received a high school diploma or general equivalency diploma (GED)
- ☐ Completed some college
- ☐ Finished college
- ☐ Finished graduate school, law school, or medical school
- ☐ Don't know

q6 **Think of the scale below as a way of showing where people stand in the United States. At the top (number 10) are the people who are <u>most well off</u> – those who have the most money, the best education, and the most respected jobs. At the bottom (number 1) are the people who are <u>least well off</u> – those who have the least money, the poorest education, and the least respected job or no job. On the scale, select the number that best represents where you think <u>you</u> stand at this time of your life, relative to other people in the United States.**

- ☐ 10 – Most well off
- ☐ 9
- ☐ 8
- ☐ 7
- ☐ 6
- ☐ 5
- ☐ 4
- ☐ 3
- ☐ 2
- ☐ 1 – Least well off

q7 **In the last 12 months, have you received any information on or learned something about any of the following topics? Select all that apply. *Please indicate information received from any type of source, for example at school or church, from health professionals, or from friends or family members.***

- ☐ Abstinence from sex or how to avoid having sex
- ☐ Methods of birth control or where to get birth control
- ☐ Condoms
- ☐ Sexually transmitted diseases or infections (STDs or STIs)
- ☐ How to talk to a partner about consent and whether or not to have sex
- ☐ How to talk to a partner about whether or not to use condoms or birth control
- ☐ How to say no to sex
- ☐ Pregnancy
- ☐ Safe sexual relationships
- ☐ I haven't received any information or learned anything about these topics in the past 12 months

q8 **Did you receive or learn about any of this information from a formal class/program or a health care provider?**

- ☐ Yes
- ☐ No

q9 **In the last 12 months, which of the following types of health services have you received? Select all that apply.**

- ☐ Mental health (for example, counseling for depression or anxiety)
- ☐ Sexual/reproductive health (for example, fertility services, sexually transmitted infection and HIV testing and treatment, or to get birth control)
- ☐ Primary health care services (for example, an annual physical examination, vaccination, or treatment for an illness)

SECTION 2: RECENT BEHAVIORS AND EXPERIENCES

Now we will ask you some questions having to do with your recent behaviors and experiences. There are no right or wrong answers. For each question, choose the answer that best represents YOUR experience. The information that you provide is very valuable and will help us understand the experiences of people your age. If you are not certain of an answer, please provide your best guess. Remember, your answers are strictly confidential. Your name is not on this questionnaire and will not be associated with any of your responses.

q10 **In the past 30 days, on how many days did you smoke cigarettes? *If you do not know the exact number, provide your best guess. If you have not smoked cigarettes in the past 30 days, your answer will be 0 days.***

_____day(s)

q11 **In the past 30 days, on how many days did you have at least one drink of alcohol? *If you do not know the exact number, provide your best guess. If you have not had any alcohol in the past 30 days, your answer will be 0 days.***

_____day(s)

q12 **Please indicate which method of prescription birth control you are currently using? Select only one.**

- ☐ Oral contraceptives (for example, the pill)
- ☐ The patch (for example, Ortho Evra)
- ☐ The shot/injection (for example, Depo-Provera)
- ☐ The ring (for example, NuvaRing)
- ☐ Implant (for example, Implanon or Nexplanon)
- ☐ IUD (for example, ParaGard, Skyla, Mirena, Kyleena, Liletta)
- ☐ None of the above

q13 **What is your main reason(s) for using prescription birth control? Select all that apply.**

- ☐ Pregnancy planning and spacing
- ☐ Pregnancy prevention

☐ Reduction in menstrual-related side effects (for example, PMS, acne, or migraines)

☐ Reduction in risk for certain types of cancers (for example, endometrial cancer or ovarian cancer)

☐ Treatment for menstrual-related symptoms (for example, severe menstrual pain or heavy bleeding)

☐ Other (please specify):

_____

TIP: In the next set of questions, when we ask about vaginal sex, we mean when a penis is put in a vagina. *Please do not report on episodes of oral or anal sex here; we will ask about these types of sex in a later section of the questionnaire.*

q14 **In the past 30 days, how many times have you had vaginal sex?** ***If you do not know the exact number, provide your best guess. If you have not had vaginal sex in the past 30 days, your answer will be 0 times.***

_____time(s)

q15 **In the past 30 days, how many times have you had vaginal sex without using a condom?** ***If you do not know the exact number, provide your best guess. If you have not used a condom during vaginal sex in the past 30 days, your answer will be X - the number of times you said you had vaginal sex in the past 30 days.***

_____time(s)

q15_alt **In the past 30 days, how many times have you had vaginal sex without using a condom?** ***If you do not know the exact number, provide your best guess.***

_____time(s)

TIP: Below, when we ask about prescription birth control, we are talking about the following methods of birth control that might be prescribed or administered to you by a health care provider (for example, a doctor or nurse practitioner).

* Oral contraceptives (for example, the pill)
* The patch (for example, Ortho Evra)
* The shot (for example, Depo Provera)
* The ring (for example, NuvaRing)
* The implant (for example, Implanon or Nexplanon)
* IUD (for example, ParaGard, Skyla, Mirena, Kyleena, or Liletta)

Remember, being "protected" means you used your birth control correctly as directed or prescribed.

q16 **In the past 30 days, how many times have you had vaginal sex without being protected by some form of prescription birth control?** ***If you do not know the exact number, provide your best guess. If you have not used prescription birth control in the past 30 days, your answer will be X - the number of times you said you had vaginal sex in the past 30 days.***

_____time(s)

q16_alt **In the past 30 days, how many times have you had vaginal sex without being protected by some form of prescription birth control.** ***If you do not know the exact number, provide your best guess.***

_____time(s)

TIP: Below, when we ask about prescription birth control, we are talking about the following methods of birth control that might be prescribed or administered to you by a health care provider (for example, a doctor or nurse practitioner).

* Oral contraceptives (for example, the pill)
* The patch (for example, Ortho Evra)
* The shot (for example, Depo Provera)
* The ring (for example, NuvaRing)
* The implant (for example, Implanon or Nexplanon)
* IUD (for example, ParaGard, Skyla, Mirena, Kyleena, or Liletta)

Remember, being "protected" means you used your birth control correctly as directed or prescribed.

q17 **In the <u>past 30 days</u>, how many <u>times</u> have you used emergency contraception (for example, ella or Plan B) after vaginal sex to prevent pregnancy?** ***If you have not used emergency contraception in the past 30 days, your response will be 0 times.***

_____time(s)

q18 **In the <u>past 30 days</u>, how many times did you have vaginal sex using <u>both</u> a condom <u>and</u> one of the listed forms of prescription birth control <u>at the same time</u>?** ***If you do not know the exact number, provide your best guess. If you used two methods of protection every time you had sex, your answer will be X – the number of times you said you had vaginal sex in the past 30 days.***

_____time(s)

q18_alt **In the <u>past 30 days</u>, how many times did you have vaginal sex using <u>both</u> a condom <u>and</u> one of the listed forms of prescription birth control <u>at the same time</u>?** ***If you do not know the exact number, provide your best guess.***

_____time(s)

q19 **Considering the <u>LAST</u> time you had <u>vaginal</u> sex, which of the following did you use? Select all that apply.**

- ☐ Oral contraceptives (for example, the pill)
- ☐ The patch (for example, Ortho Evra)
- ☐ The shot/injection (for example, Depo-Provera)
- ☐ The ring (for example, NuvaRing)
- ☐ Implant (for example, Implanon or Nexplanon)
- ☐ IUD (for example, ParaGard, Skyla, Mirena, Kyleena, Liletta)
- ☐ Emergency contraception (ella, Plan B)
- ☐ Condoms
- ☐ The sponge
- ☐ Diaphragm
- ☐ Foam or spermicide
- ☐ Natural family planning (rhythm method)
- ☐ I did not use any of these

TIP: Below, when we ask about oral sex, we mean when one person puts their mouth in contact with another person's genitals – meaning the penis, vagina, or anus. You are considered to have had oral sex regardless as to whether you "gave" or "received" it.

q20 **In the past 30 days, how many times have you had oral sex?** ***If you do not know the exact number, provide your best guess. If you have not had oral sex in the past 30 days, your response will be 0 times.***

_____time(s)

TIP: Below, when we ask about anal sex, we mean when a penis is put into another person's anus, or butt.

q21 **In the past 30 days, how many times have you had anal sex?** ***If you do not know the exact number, provide your best guess. If you have not had anal sex in the past 30 days, your response will be 0 times.***

_____time(s)

q22 **In the past 30 days, how many times have you had anal sex without using a condom?** ***If you do not know the exact number, provide your best guess. If you have not used a condom during anal sex in the past 30 days, your answer will be X – the number of times you said you had anal sex in the past 30 days.***

_____time(s)

q22_alt **In the past 30 days, how many times have you had anal sex without using a condom?** ***If you do not know the exact number, provide your best guess.***

_____time(s)

q23 **In the past 30 days, how many times did you use alcohol or drugs before having any type of sex (vaginal, oral, or anal)?** ***If you do not know the exact number, provide your best guess. If you have not used alcohol or drugs before sex in the past 30 days, your response will be 0 times.***

_____time(s)

q24 **In the past 12 months, have you conducted a self-breast exam?**

- ☐ Yes
- ☐ No

q25 **Have you ever had a conversation with a family member about your biologic family's health history?** ***When we say family health history, we mean the diseases and health conditions experienced by any member of your biologic family with whom you are genetically connected.***

- ☐ Yes
- ☐ No

TIP: Below, when we say sexual partner, we mean anyone with whom you have had anal, oral, or vaginal sex.

q26 **How many sexual partners have you had in the past 3 months?** ***If you have not had any type of sex (vaginal, oral, or anal) in the past 3 months, your answer will be 0. If you do not know the exact number, provide your best guess.***

_____ partner(s)

q27 **Thinking of your interaction(s) with your sexual partner(s) in the past 3 months, please indicate how many sexual partner(s) you have had in each category:**

Serious relationship: _____ partner(s)

Casually dating, but not serious: _____ partner(s)

Sleeping with, but not dating: _____ partner(s)

One-night stand: _____ partner(s)

SECTION 3: CONDOM AND CONTRACEPTIVE KNOWLEDGE

In this section, we ask about different methods people use to prevent pregnancy and sexually transmitted infections or diseases (STIs or STDs). We are interested in your personal knowledge of these different methods of protection. Remember, your name is not on this questionnaire, and your answers are confidential.

q28 **Please indicate whether the following statements are True or False. If you are not certain of the correct answer, please select Don't know.**

	True or False?		
	True	False	Don't know
Condoms have an expiration date.	☐	☐	☐
Some methods of emergency contraception (such as ella or Plan B) require a prescription.	☐	☐	☐
Birth control pills are effective even if someone misses taking them for two or three days in a row.	☐	☐	☐
Condoms work well to prevent sexually transmitted infections (STIs).	☐	☐	☐
Long-acting methods of contraception like the implant or IUD cannot be removed early, even if someone changes their mind about wanting to get pregnant.	☐	☐	☐
Plan B and ella are pills that can be taken shortly after having unprotected sex to prevent pregnancy.	☐	☐	☐
The birth control pill, ring, and patch are just as effective at preventing pregnancy as IUDs and the implant.	☐	☐	☐
Condoms are not as effective at preventing pregnancy as prescription birth control methods (for example, the pill, the patch, the ring, the shot, IUD, or implant).	☐	☐	☐

SECTION 4: EMOTIONS AND FEELINGS

In this section, we want to ask you about your emotions and feelings. For each question, choose the answer that best represents how YOU feel or what YOU think. Some of these questions ask about your emotions and feelings related to sex. If you are not having sex nor intending to have sex, please answer how you think you WOULD feel if you were having sex.

TIP: As a reminder, below, when we ask about prescription birth control, we are talking about the following methods of birth control that might be prescribed or administered to you by a health care provider (for example, a doctor or nurse practitioner).

* Oral contraceptives (for example, the pill)
* The patch (for example, Ortho Evra)
* The shot (for example, Depo Provera)
* The ring (for example, NuvaRing)
* The implant (for example, Implanon or Nexplanon)
* IUD (for example, ParaGard, Skyla, Mirena, Kyleena, or Liletta)

q29 **Please rate on a scale from 1 to 7 how confident or sure you are that you could do each of the things described. The higher the number you select, the more confident you are that you could do it. *An answer of 1 means you are not at all confident that you could do what the statement is describing; an answer of 7 means you are extremely confident that you could do what the statement is***

	How confident are you?						
	(1) Not at all confident	(2)	(3)	(4) Somewhat confident	(5)	(6)	(7) Extremely confident
Tell someone you plan to have sex with that you want to use condoms?	☐	☐	☐	☐	☐	☐	☐
Convince a partner to use condoms, even if you are using some other kind of birth control (for example, the pill)?	☐	☐	☐	☐	☐	☐	☐
Insist that a condom be used?	☐	☐	☐	☐	☐	☐	☐
Refuse to have sex if a partner won't use a condom?	☐	☐	☐	☐	☐	☐	☐

Tell a partner that you do not want to have sex?	☐	☐	☐	☐	☐	☐	☐
Use a condom every time that you have sex?	☐	☐	☐	☐	☐	☐	☐
Use a condom <u>correctly</u> every time you have sex?	☐	☐	☐	☐	☐	☐	☐
Use a condom after you have been drinking?	☐	☐	☐	☐	☐	☐	☐
Use prescription birth control as directed?	☐	☐	☐	☐	☐	☐	☐
Resist having sex if you are not using some form of prescription birth control?	☐	☐	☐	☐	☐	☐	☐
Make unpleasant thoughts go away?	☐	☐	☐	☐	☐	☐	☐
Take your mind off unpleasant thoughts?	☐	☐	☐	☐	☐	☐	☐
Stop yourself from being upset by unpleasant thoughts?	☐	☐	☐	☐	☐	☐	☐
Keep from feeling sad?	☐	☐	☐	☐	☐	☐	☐

q31 **Are you currently pregnant?**

- ☐ Yes
- ☐ No
- ☐ Not sure

q32 **Thinking about your life right now, how important is it to you to avoid becoming pregnant?**

- Very important
- Somewhat important
- A little important
- Not important

q33 **If you found out today that you were pregnant, how would you feel?**

- Very upset
- A little upset
- A little pleased
- Very pleased

SECTION 5: QUESTIONNAIRE EXPERIENCE

In this final section, we ask a few questions about your experience while completing this questionnaire. When you are finished, please make sure to click the SUBMIT button at the bottom of the page.

q34 **Have you been as honest as possible in responding to all of the questions in this questionnaire?**

- ☐ Yes, all of the time
- ☐ Yes, almost all of the time
- ☐ Yes, but just some of the time
- ☐ No, none of the time

q35 **Did you have enough privacy when completing the questionnaire?**

- ☐ Yes
- ☐ No

q36 **Were there any disruptions while you completed the questionnaire?**

- ☐ Yes (please explain):_____
- ☐ No