Safer Sex Intervention: Final Impact Report

Teen Pregnancy Prevention Replication Study

Contract # HHSP23320095624WC
Order # HHSP23337011T

October 2018

Prepared for:
Amy Farb
Office of Adolescent Health
Lisa Trivits
Sarah Oberlander
Office of the Assistant Secretary for Planning and Evaluation
U.S. Department of Health and Human Services

Submitted by:

Meredith Kelsey
Jean Layzer
Cristofer Price
Kimberly Francis

Abt Associates
10 Fawcett Street
Cambridge, MA 02138

In Association with:
Decision Information Resources, Inc.
and
Belmont Research Associates

Recommended Citation

Kelsey, M., Layzer, J., Price, C., and Francis, K. “Safer Sex Intervention Final Impact Report: Findings from the Teen Pregnancy Prevention Replication Study”. Cambridge, MA: Abt Associates Inc., October 2018.

This document is available at https://aspe.hhs.gov/teen-pregnancy-prevention-tpp-replication-study

This document was prepared under contract number HHSP23320095624WC Order No. HHSP23337011T (awarded in September 2011) from the U.S. Department of Health and Human Services, Office of Adolescent Health and the Office of the Assistant Secretary for Planning and Evaluation. Any statements expressed are those of the authors and do not necessarily reflect the views of the Office of Adolescent Health, the Office of the Assistant Secretary for Planning and Evaluation, or the U.S. Department of Health and Human Services.

Acknowledgements

We would like to thank many individuals and organizations who helped to make this evaluation possible. First, we thank the staff from the Office of Adolescent Health (OAH) within the U.S. Department of Health and Human Services (HHS) for their support throughout the project. In particular Amy Farb, the OAH project officer for the Teen Pregnancy Prevention Replication Study, provided critical guidance throughout the project. We also thank the many program officers and Alexandra Warner. Lisa Trivits and Sarah Oberlander, the project officers for the TPP Replication Study from the Office of the Assistant Secretary for Planning and Evaluation (ASPE) provided critical direction and support throughout the project and helped to ensure the success of the study.

We would like to thank our project partners Decision Information Resources (DIR) for their commitment and hard work in support of the study and data collection efforts, along with Innovative Clinical Research Solutions (ICRS) and CiviCore. Belmont Research Associates provided invaluable leadership and project direction, and helped navigate many obstacles and establish critical relationships on the ground. We would like to acknowledge each of the grantee sites whose close partnership and commitment were essential to the success of the study. Project staff not only implemented the program, but participated in interviews and remained dedicated to the study throughout. Thanks to all of the health educators and project directors for sharing their insights about motivational interviewing in a clinic context – the contribution to the field would not be possible without them.

At Abt, the project was supported by a team of individuals who worked closely with each of the grantee sites. Jessica Walker, Kim Francis and Michelle Martin provided essential support for implementation of the study. Rob Olsen and Randall Juras provided guidance and support during the initial study design phase, and Beth Boulay provided valuable feedback on the report. Thanks to Dan Weiss, Heather He and Jessica Walker for excellent programming assistance, and Stefanie Falzone and Bry Pollack for producing and editing the report.

And finally, a tremendous thanks to all of the young women who participated in the study. Without them, this entire endeavor would not have been possible.

Contents

Exhibits

1. Introduction

Reducing rates of unplanned teen pregnancy and sexually transmitted infections (STIs) is a priority for the U.S. Department of Health and Human Services (HHS). To achieve this goal, the Department is investing in evidence-based pregnancy reduction strategies and targeting populations at highest risk for teen pregnancy. The federal Teen Pregnancy Prevention (TPP) Program, administered by the Office of Adolescent Health (OAH), includes funding for programs that are intended to address high rates of teenage pregnancy by (1) replicating evidence-based models, and (2) testing innovative strategies.

The TPP Program was authorized in 2010 as part of the larger Teen Pregnancy Prevention Initiative and initially included $100 million in annual funding to support programming. Of these funds, $75 million were available annually to support five-year grants for replicating 28 program models that prior rigorous evaluations had shown to be effective. These program models were identified through a systematic, comprehensive review of the literature on prevention of teen pregnancy, STIs, and sexual risk behaviors (Kappeler & Farb, 2014).

The TPP Program acknowledges the limitations of existing research and the need for additional research on programs, citing lessons learned from the comprehensive evidence review, such as an absence of independent evaluations and a limited number of program replications (Goesling et al., 2014). The review highlighted that the evidence for many of the 28 programs eligible for replication rested on single studies of effectiveness, often conducted a long time ago and with a single population. A program may work in one location with a particular population, but that does not necessarily mean it will be effective in another. Further, implementing a program model with fidelity often competes with the need to adapt to local conditions. For these reasons, a study of multiple replications of selected program models is an important contribution to the existing research.

1.1 The TPP Replication Study

The TPP Replication Study[1] was conducted for HHS, under a contract with the Office of the Assistant Secretary for Planning and Evaluation (ASPE) and OAH, by Abt Associates and its subcontractors Belmont Research Associates, Decision Information Resources (DIR), and CiviCore. The study has two major components: an impact study and an implementation study.

Impact Study. Through a series of rigorous experimental design evaluations, the impact study tests multiple replications of three evidence-based program models to determine their effectiveness across different settings and populations.

Implementation Study. A comprehensive implementation study provides information about the contexts in which the evidence-based programs were implemented, the challenges faced in implementing them, and aspects of program implementation that help to explain program impacts.

1.2 The Three Models Replicated

ASPE and OAH selected three program models from the initial cohort of TPP-funded grants to test and replicate:

[1] The study was also referred to as the Teen Health Empowerment Study in the field with program staff and study participants.

- *Safer Sex Intervention*, a clinic-based HIV/STI prevention program for high-risk adolescent females;
- *Reducing the Risk*, a sexual health education curriculum; and
- *¡Cuídate!* an HIV/STI risk reduction program targeting Latino youth.

Criteria used in the selection of these models included the breadth and scale of the proposed replication effort and the number of grantees that proposed to replicate a model.[2] In addition, the three models represent a range of targeting and service strategies, as well as some variation in the settings in which services are provided.

1.3 Focus of This Report

This report focuses on the *Safer Sex Intervention (SSI)*, presenting findings from two follow-up surveys designed to examine its short-term and longer-term impacts. It is one in a series of reports that present findings on the implementation and effectiveness of the three program models. Three implementation study reports document the implementation of each of the three models. In addition, nine site profiles provide an overview of the program implementation, as well as descriptive information about the study participants at baseline in each site.[3]

[2] The *Teen Outreach Program (TOP)* was the most-frequently replicated program model. There were seven independent evaluations as a condition of the grants. For this reason, it was excluded from consideration for the TPP Replication Study. *Becoming a Responsible Teen (BART)*, another widely-used model, was also excluded because it had already undergone several evaluations. All three models selected were originally proposed by at least five grantees.

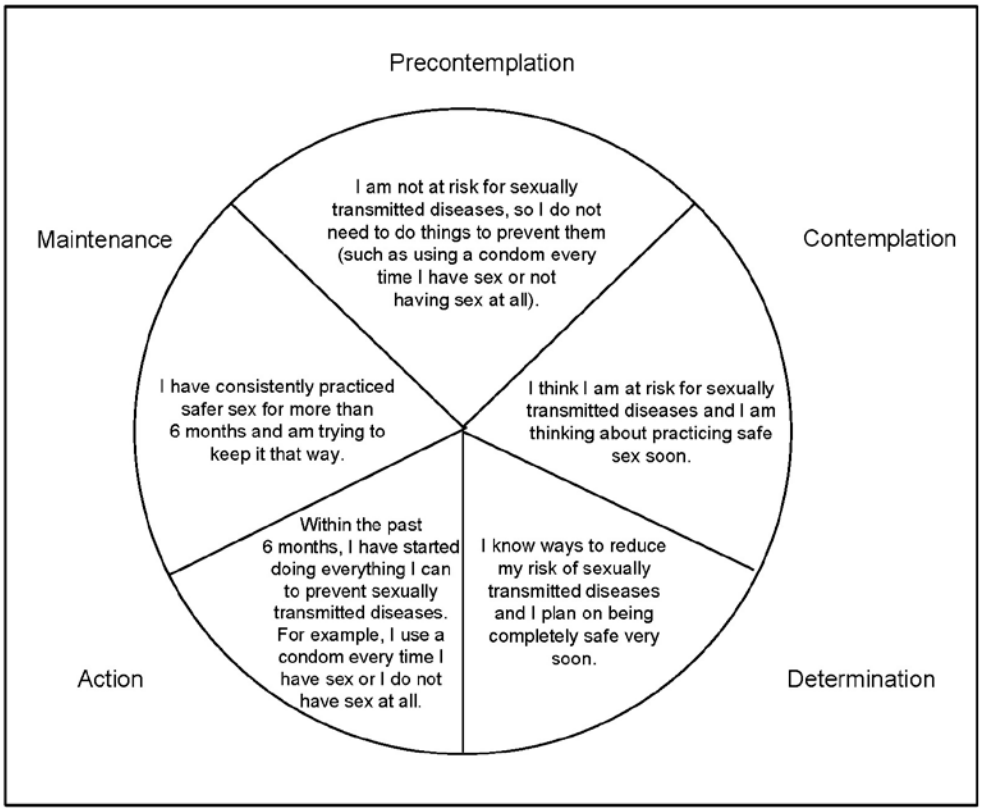
[3] The profiles are available at: https://aspe.hhs.gov/basic-report/tpp-replication-study

2. The Program Model: *Safer Sex Intervention*

The *Safer Sex Intervention (SSI)* is a clinic-based intervention intended to reduce the incidence of STIs and increase condom use among higher-risk, sexually active female adolescents. The intervention is delivered in one-on-one, face-to-face sessions with a female health educator. It has two versions: the Pre-Contemplation Stage Module, which emphasizes delivering information and obtaining feedback about safer sex behaviors; and the Contemplation Stage Module, which emphasizes education, skills, self-efficacy, and self-esteem. The choice of which version to use is made by the health educator on the basis of the client's self-assessment on the Wheel of Change tool (Exhibit 2.1), their subsequent discussion, and the health educator's own assessment of the client.

Using a videotape to introduce information about condom use, the Wheel of Change for self-assessment and reflection, and a motivational interviewing strategy to encourage participant-directed discussion, the health educator guides the client through a sequence of topics and allows time for role plays, questions, and feedback on the session. Intervention topics include the consequences of unprotected sex, risk perception, preventing pregnancy and STIs, condoms, where to obtain condoms, secondary abstinence, and talking about sex (Exhibit 2.2). After the initial 50- to 60-minute session, three subsequent booster sessions, similar in content, are delivered one, three, and six months later. These booster sessions can vary in length from 10 to 20 minutes, depending on the needs of the client. The health educator uses them to review information, assess progress, and provide the client with additional information and practice, if needed. Clients are offered condoms and informational materials.

Exhibit 2.1: *Safer Sex Intervention*'s Wheel of Change Tool



Source: Shrier et al., 2001.

Exhibit 2.2: *Safer Sex Intervention*, Core Elements and Topics

Core Element	Topics/Activities
Introduction and overview	• Introductions and discussion of SSI goals
Stage of change determination	• Wheel of Change explanation • Wheel of Change stage chosen
Consequences of unprotected sex	• Elicit examples of consequences of unprotected sex • Review STI facts • Female anatomical model used to discuss STI risk to females and demonstrate the ascension of infection
Risk perception	• Discuss participant's personal risk of STI • Discuss symptoms of STIs and importance of protection every time • Elicit change talk around STI risk
Preventing the consequences	• STI/pregnancy prevention activity • "Birth Control Choices" brochure
About condoms	• Discuss participant's use of condoms • "Condoms: How to Use Them" brochure • Male condom review and condom demonstration • Female condom demonstration • Condom keychain
Obtaining condoms	• Discuss with participant where to obtain condoms • Elicit motivation to obtain condoms
Secondary abstinence	• Engage in discussion about not having sex and assess interest/motivation from participant • Brochures
Talking about sex	• Discussion about talking with your partner • Brochures
Role play	• For contemplation stage only

Source: Firpo-Triplett, Rex, & Shrier (2011).

2.1 *Safer Sex Intervention* Logic Model

The theoretical framework for *SSI* draws on *social cognitive theory*, the *transtheoretical model of behavior change*, and the technique of *motivational interviewing* (Bandura, 1986; Prochaska & DiClemente, 2005; Prochaska & Velicer, 1997). The two theories underpin many other program models in the field of pregnancy prevention, sexual health education, and beyond, influencing the content and activities of the intervention and stressing the dynamic nature of behavior change. Motivational interviewing, however, is relatively uncommon in this field, although it is widely used in other fields and with adolescents.

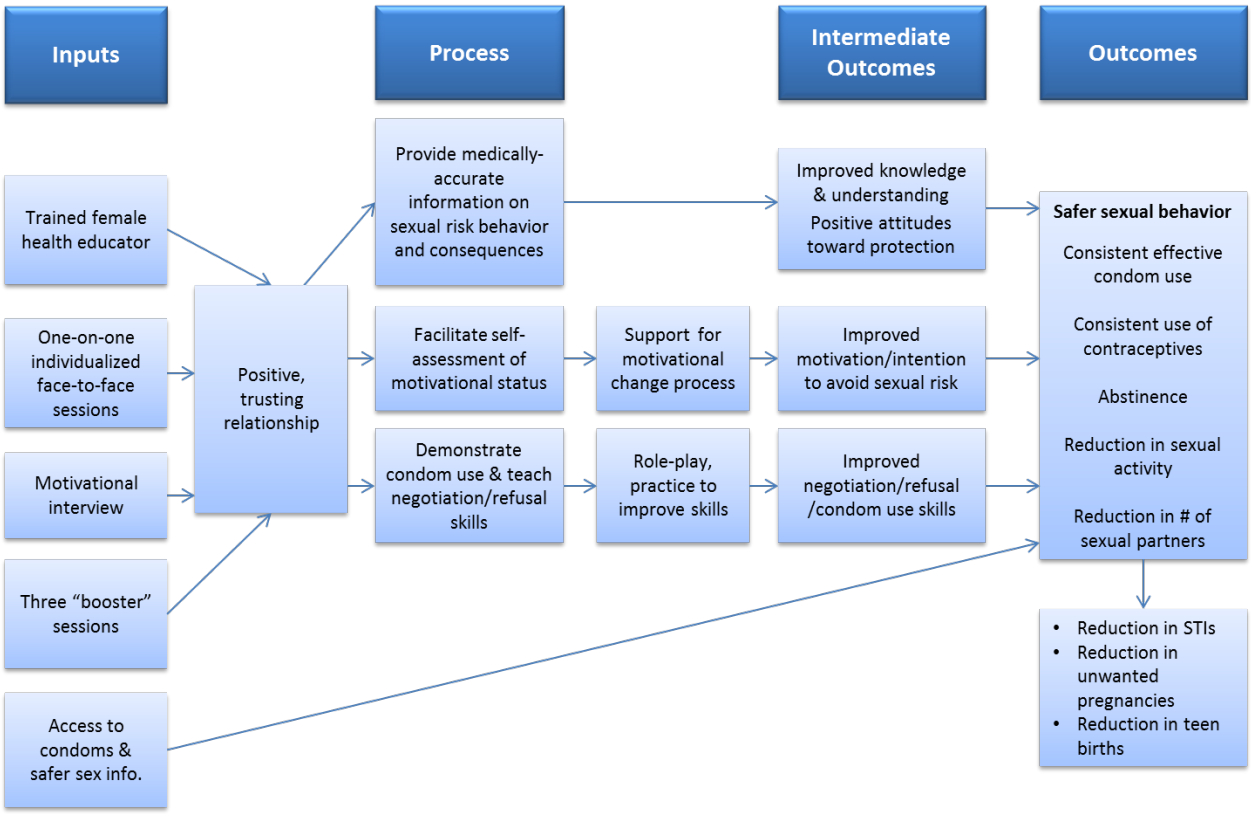
Essential to the program's strategy is the recognition that behavior change must be initiated and maintained in the face of barriers that may be unique to an individual. The role of the health educator is not that of teacher or clinician, but guide and facilitator. The *SSI* strategy allows for personalized counseling that captures the participant's attention and takes into account individual needs and challenges. During the initial session, the health educator helps the adolescent identify her needs, motivations, and

intentions; gradually identify obstacles to behavior change; and make plans to address them. Through subsequent booster sessions, the health educator tracks the participant's progress through the stages of change, from Precontemplation through Maintenance (see Exhibit 2-1).

Exhibit 2.3 shows the *SSI* program elements, its hypothesized outcomes, and the pathways by which *SSI* seeks to achieve these outcomes. The program's theory of action suggests that a trained health educator, using motivational interviewing techniques during an initial individualized counseling session and subsequent booster sessions, will establish a positive and trusting relationship with the client. In this context, the educator provides medically accurate information, facilitates self-assessment, encourages a client-directed discussion about risky sexual behavior and relationship issues, demonstrates condom use, and teaches negotiation skills.

Through question and answer, discussion, role play, and the educator's support for behavioral change, the client is expected to show improved knowledge and understanding of sexual risk behavior and its consequences, become more motivated to avoid risk, and become more able to negotiate safe sex and refuse unwanted sex. Greater understanding of the consequences of risky sexual behavior, improved motivation to avoid risk, and better negotiation skills are intermediate outcomes that are expected to lead to the outcomes of interest: safer sexual behaviors such as consistent and effective use of condoms and other contraceptives, abstaining from or reducing sexual activity, and reducing the number of sexual partners. Ultimately, these safer sexual behaviors are expected to reduce rates of STIs and unplanned pregnancies and births among teens.

Exhibit 2.3: *Safer Sex Intervention* Logic Model



In Section 3.5.2, we describe in more detail the modifications to the program model proposed by each of the organizations replicating it. Their modifications updated materials and improved retention while adhering to the core components of the model.

2.2 Prior Evidence of Effectiveness

SSI is one of two clinic-based programs identified as evidence-based by the HHS Pregnancy Prevention Evidence Review that TPP grantees could choose to implement (HHS, 2010). As with many other program models identified through OAH's comprehensive review (Kappeler & Farb, 2014) prior to this evaluation, evidence for *SSI*'s effectiveness comes from a single study by the program developer (Shrier et al., 2001) that was completed almost 20 years ago.

The *SSI* program was developed in response to high rates of STIs among high-risk adolescent girls. The program developer originally tested the intervention in an urban children's hospital adolescent clinic and inpatient service with female adolescents who presented for treatment of cervicitis or were admitted for the management of pelvic inflammatory disease (Shrier et al., 2001). Findings from that randomized controlled trial suggested that after six months, which coincided with the end of medical treatment and the program's six-month booster session with the health educator, *SSI* participants were significantly less likely than were study members who did not participate in the intervention to report having multiple sexual partners in that timeframe.

There were no other significant findings on behavioral outcomes. However, the study authors noted the suggestion of a positive effect on condom use at the six-month data collection. The study also examined knowledge of sexual risk and attitudes toward condom use and found a positive program effect on both knowledge of STI risk and positive attitudes toward condoms after one month (the interval for the first *SSI* booster session).

3. Evaluation Design and Implementation

The impact study is designed to estimate the effects of *SSI* on sexual risk behaviors and consequences, as well as on the non-behavioral, intermediate outcomes the logic model predicts will lead to the behavioral outcomes that *SSI* seeks to achieve.

In the first part of this chapter, we set forth the study's research questions and describe the design elements of the study, including the overall evaluation strategy; the measures selected to address the research questions and the timing of measurements; and the analytic strategy devised to assess program effectiveness. In the second part of the chapter, we describe our implementation of the study design and analysis plan in each of the three replication sites.

3.1 Research Questions

The following research questions guided the evaluation:

1. Did *SSI* have an impact on sexual behavior after nine months and 18 months?
2. Did *SSI* reduce the incidence of unplanned teen pregnancies after 18 months?
3. Did *SSI* reduce the incidence of STIs after 18 months?
4. Did *SSI* have an effect on non-behavioral, intermediate outcomes hypothesized to lead to behavior change (i.e., knowledge, attitudes, motivation, intentions, and skills) after nine months and 18 months?
5. Do program effects on behavior differ by replication site and for key subgroups (e.g., age, race/ethnicity)?

These five research questions imply a wide range of outcomes, including non-behavioral (intermediate) outcomes that the program model suggests are precursors of the behavioral outcomes, and the behavioral consequences that are the ultimate targets of the program and the TPP Initiative. The fifth research question is intended to take maximum advantage of pooled data from all three replications by exploring potential differences in effect for specific sites and subgroups. We elected to investigate non-behavioral and behavioral outcomes to trace the pathways of influence in the program logic model.

Collecting data and estimating effects on so many outcomes does, however, pose challenges for the ways in which data are analyzed and how the results are interpreted. The sheer number of statistical tests of effectiveness needed to address those questions means that we would expect some share of them (perhaps 5 percent) to generate statistically significant results simply by chance. In a later section of this chapter, we describe the steps we took to minimize the risk of incorrectly concluding that *SSI* had an impact.

3.2 Key Design Features

The design of the evaluation of SSI included the following key elements:

- Multiple replications of the program model (three sites);
- Within each replication site, implementation of a rigorous experimental design in which young women were randomly assigned to receive either the intervention or the usual services offered in the clinic;

- Measures that allow us to address all of the research questions;
- A measurement schedule that captures both short-term and longer-term outcomes;
- An analytic strategy that pools data from all replications to measure sexual behavior and the consequences of sexual risk behavior and to examine differences in program effectiveness by replication site, as well as for important youth subgroups; and
- A strategy that identifies a key set of behavioral outcomes and prioritizes a limited number of "confirmatory" analyses to increase confidence in the study findings. At the same time, the strategy also allows for "exploratory" (and more speculative) analyses that incorporate many more outcomes, both behavioral and non-behavioral.

3.3 Measures and Measurement Schedule

Outcome measures selected for the study fall into three categories: sexual activity and sexual risk behavior, sexual consequences, and non-behavioral intermediate outcomes. Exhibit 3.1 summarizes the outcome measures and their construction; Appendix A provides a more complete description of each measure and its individual items.

Exhibit 3.1: Outcome Measures

Measure	Definition
Sexual Behavior Outcomes	
Sexual activity	
Recent sexual activity(in last 90 days) [a] Sexual intercourse in the last 90 days Oral sex in the last 90 days Anal sex in the last 90 days	Single items, scored 1 (yes) or 0 (no)
Sexual risk behavior	
Sexual intercourse without any birth control (in last 90 days)[a] Sexual intercourse without a condom (in last 90 days) Oral sex without a condom (in last 90 days) Anal sex without a condom (in last 90 days)	Single items, scored 1 (yes) or 0 (no)
Sexual intercourse with more than one partner (lifetime) Sexual intercourse with more than five partners (lifetime)	Single items, scored 0 or 1, with 1 representing multiple sexual partners in one's lifetime
Sexual Consequences (Longer-term follow-up only)	
Pregnant since baseline[a] Diagnosed with STI in the last 12 months	Single items, scored 1 (yes) or 0 (no)
Non-Behavioral Intermediate Outcomes	
Knowledge	
Knowledge of pregnancy risk	Continuous index: Average of responses to four questions about circumstances in which it is possible to become pregnant and the extent to which contraceptive methods protect against pregnancy, multiplied by 100. Average scores range from 0 to 100 and represent the percentage of the four questions answered correctly, with higher values representing more accurate knowledge.
Knowledge of STI risk	Continuous index: Average of responses to 12 questions about STI transmission and prevention, multiplied by 100. Average scores range from 0 to 100 and represent the percentage of the 12 questions answered correctly, with higher values representing more accurate knowledge.
Attitudes	
Attitudes toward protection	Continuous index: Average of responses to 12 questions about attitudes toward using condoms and/or birth control during sex. Average scores range from 1 to 4, with higher values representing more positive attitudes toward using protection.

Measure	Definition
Attitudes toward risky sexual behavior	Continuous index: Average score of seven binary items about the acceptability of risky sexual behavior, multiplied by 100 to represent the percentage of items agreed with. Average scores range from 0 to 100, with higher values representing more support for risky behavior.
Motivation	
Motivation to delay childbearing	Continuous index: Average of three items about motivation to delay childbearing. Scores range from 1 to 4, with higher values representing greater levels of motivation.
Intentions (in next 12 months)	
Intention to have sexual intercourse Intention to have oral sex Intention use a condom if having sexual intercourse Intention to use birth control if having sexual intercourse	Single items, scored 0 or 1, with 1 representing stronger intention.
Skills	
Refusal skills	Continuous index: Average of responses to six questions about perceived ability to refuse to engage in risky sexual behavior. Scores range from 1 to 4, with higher values representing greater certainty about refusal skills.
Condom negotiation skills	Continuous index: Average of responses to seven questions about perceived ability to obtain and negotiate the use of condoms. Scores range from 1 to 4, with higher values representing greater certainty about condom negotiation skills.

[a] Designated as a key outcome for confirmatory analyses (see Section 3.4.2).

The study design called for young women in the three replication sites to be surveyed three times: before the intervention began (baseline); nine months after the baseline survey (short-term follow-up); and 18 months after the baseline survey (longer-term follow-up). This schedule allowed us to capture behavioral outcomes that we might expect to occur immediately, as well as the consequences of sexual behaviors that may not occur until the longer term (such as pregnancy).

3.4 Analytic Approach

Two strategic decisions shaped the analysis of data collected over the life of the study. The first was a decision about how to treat the three replications of the program. The second was a decision about prioritizing analyses to answer the key research questions. Each of these decisions as they relate to our analytic approach is described below.

3.4.1 Incorporating Three Program Replications

When deciding how to treat the three replications of *SSI*, one possibility was to treat them as three stand-alone evaluations; Abt staff designed each of the three evaluations independently, taking into account any special circumstances in each replication site (e.g., at the grantee's request, surveys of youth in the Knox County, Tennessee, replication excluded questions about anal sex). The sample requirements in each of the replication sites were calculated to permit detection of relatively small impacts on sexual behavior.

The other possibility, the one ultimately selected, was to consider the three evaluations as components of an integrated study, in which data were pooled across the three sites. This strategy offered several benefits. Importantly, the tripled sample size would allow us to estimate the impact of *SSI* on likely consequences of sexual risk behavior, such as pregnancy and diagnosis of an STI. Prevention of these consequences is the primary goal of the TPP Initiative, but measuring them as part of an evaluation is a challenge. Given that these outcomes are relatively rare events, the sample size necessary to detect a possible intervention impact on pregnancies and STIs requires resources beyond what is available in many single-site studies.

In addition, pooling data across the three replication sites would allow us to conduct the many subgroup analyses necessary to address the study's research questions. Subgroup analyses would be less feasible with the smaller sample sizes of the individual replications. Even with pooling the data across sites, we also have the ability to examine the extent to which replications differed in their effectiveness.

Finally, although three replications cannot be held to represent the universe of possible replications, findings from the analysis of pooled data would have greater generalizability than findings from any single-replication study. An integrated study would include a variety of settings, a range of ages, and variation in other demographic characteristics.

A decision to create an integrated evaluation in which data from all three replications would be pooled for analytic purposes was supported by OAH's requirements of grantees to define, measure, and adhere to fidelity to the program model. These requirements ensured that each of the three replications implemented the same core program elements. The random assignment, measurement, and data collection procedures described elsewhere in this chapter were also the same across the replication sites. The consistency of these design elements ensured that impact estimates derived from data pooled at the program level would represent rigorous tests of a well-defined and well-implemented program model.

For all these reasons, we elected to pool the data from all three replication sites.

3.4.2 Prioritizing the Analyses Needed to Answer Key Research Questions

We noted earlier that the study's research questions demonstrate interest in a variety of outcomes, both behavioral and non-behavioral, as well as interest in understanding the extent to which the program works differently for different replication sites and different subgroups. In practical terms, exploring these multiple interests translates into a large number of statistical tests, some of which will produce statistically significant impact findings simply by chance. To reduce the risk of spurious findings, we needed to develop a strategy that assigned the greatest weight to analyses of greatest interest to federal policymakers.

The first step was to identify a small set of behavioral outcomes by which the success of the *SSI* program would be judged. These outcomes reflect the goals of the federal TPP Program and of most of the interventions funded by it. These outcomes span both short- and longer-term measurement points. Exhibit 3.2 shows the measurement domains and the key outcomes we identified.

Exhibit 3.2: Measurement Domains and Key Outcomes

Measurement Domain	Outcomes
Recent sexual behavior at the short-term follow-up	1. Sexual activity in the last 90 days 2. Sexual intercourse without birth control in the last 90 days
Recent sexual behavior at the longer-term follow-up	1. Sexual activity in the last 90 days 2. Sexual intercourse without birth control in the last 90 days
Consequences of sexual risk behavior	1. Pregnancy since baseline[a]

[a] The pregnancy outcome was reported only at the longer-term follow-up because of the low prevalence rate and statistical considerations.

The second step was to identify the sample on which to test impacts of *SSI* on these key outcomes. Given the advantages of a large, diverse sample, we selected the full sample, pooling data across the three replication sites as discussed in Section 3.4.1.

We use findings from the analyses of these prioritized outcomes ("confirmatory analyses") to make claims about the impact of *SSI*. The confirmatory analyses estimate impacts of *SSI* on the key outcomes for the full sample, using data pooled across the three replication sites. Additional analyses, testing

different outcomes or using different samples or subgroups ("exploratory analyses"), should be interpreted as suggestive of potential effects (see Schochet, 2008a).

In the last section of this chapter, we describe in more detail how the impact analyses were conducted and the procedures for making statistical corrections for multiple comparisons.

3.5 Implementing the Study Design

This section describes the selection of the three replication sites, site-specific program designs, settings for the program, the treatment and control conditions, recruitment and random assignment, and our data collections strategy.

3.5.1 Selection of Replication Grantees

The study design called for evaluating at least three replications of *SSI*, which, at the time of site selection for the study, was being replicated by five grantees. Complicating site recruitment was that most grantees had not planned for a rigorous evaluation.[4] One of the five grantees was eliminated because of concerns it would not be able to build a sufficient sample of youth in two years, the period estimated to achieve the required study sample size. A second grantee was eliminated due to concerns about sample size combined with other considerations that could impede a strong test of the model, leaving three of the five potential candidates.

The three grantees selected are described below:

- **Hennepin County Human Services and Public Health Department (Minnesota)** has long played a leadership role in serving at-risk youth and ensuring the health and well-being of youth and families. For more than 30 years, the department has provided programming and research support for early childhood education, improving high school graduation rates, and preventing adolescent drug and alcohol use. It has partnered with various community agencies to deliver evidence-based programs and provide teen pregnancy prevention services.
- **Knox County Health Department (Tennessee)** is the local public health agency serving the City of Knoxville and Knox County. Its Community Assessment and Health Promotion unit, with nine full-time health educators, provides primary prevention services in the areas of adolescent pregnancy, sexually transmitted diseases, sexual violence, injury, child safety and childhood diseases.
- **Planned Parenthood of Greater Orlando (Florida)**, an affiliate of Planned Parenthood Federation of America, Inc., operates as a community-based non-profit 501(c)(3) organization.[5] Since 1995, PPGO has provided reproductive health services (on a fee-for-service basis) and sexual health education in four central Florida counties—Orange, Osceola, Seminole, and Brevard.

---

[4] The 2010 TPP grant program included multiple funding ranges. All funded projects were expected to monitor and report on program implementation and outcomes through performance measures. Projects in the higher funding ranges (greater than $1 million per year) were expected to be implemented in multiple sites within a targeted geographic area and were required to have an independent local evaluation. Two of the *SSI* replications selected for the study were in the lower funding range (less than $1 million per year) and so were not expected to have a rigorous local evaluation. Hennepin County, a larger-scale replication, had proposed a rigorous local evaluation.

[5] In July 2015, PPGO merged with another Planned Parenthood affiliate to become Planned Parenthood of Southwest and Central Florida.

3.5.2 Site-Specific Program Designs

In all three replication sites, *SSI* grantees served young women aged 13 – 19 who were sexually active or about to become sexually active and not pregnant or parenting at the time of enrollment. This is a broader population than the original intervention (Shrier et al., 2001), which targeted youth who had just been diagnosed with an STI. This change in target population was proposed at the outset by the grantees, with the exception of PPGO, which had proposed to replicate the intervention with young women coming into the clinic for STI screening (close to the population of the original study).[6] During the pilot year, PPGO requested approval to serve a broader population of sexually active (or about to become sexually active) young women. OAH and *SSI*'s developer approved this adaptation for all three *SSI* grantees.[7] Other approved adaptations that were implemented in all three replication sites included replacing the original video, which was outdated, as recommended by the developer.[8] PPGO and Hennepin also successfully implemented an approved adaptation that enabled educators to conduct booster sessions remotely via video chat (e.g., Skype or FaceTime) instead of in the clinic.

Each of the replications was required to implement the program with fidelity to the *SSI* model, and fidelity was assessed, monitored, and reported to OAH at regular intervals by program staff. OAH required all of its TPP Program grantees to observe 10 percent of sessions to monitor program implementation quality and fidelity. However, given the individualized nature of the intervention and the heavy reliance on the establishment of a personal rapport and trusting relationship between the health educators and young women, the OAH requirement for observations of sessions was waived for the grantees implementing *SSI*.

3.5.3 Settings for the Program

SSI was implemented in clinics within each of the replication sites. In each of the clinics, the intervention was considered a separate educational offering. Health educators were given office space/exam rooms within the clinic, and clinicians identified and referred eligible young women to the program. The extent to which the intervention was integrated into the standard set of clinic services varied across replication sites and clinics.

Hennepin County Human Services and Public Health Department, the largest replication site, offered *SSI* in 19 different clinics during the study enrollment period. It contracted with provider agencies to deliver *SSI* to at-risk youth from various racial and ethnic backgrounds in areas with the highest teen birth rates in the county. The clinics included seven school-based clinics, five community-based clinics, four teen health clinics, one hospital-based pediatric clinic, one STI/public health clinic, and one clinic for

---

6 Hennepin County had originally proposed serving males as well as females, but this adaptation was not approved by OAH because there was no prior evidence of effectiveness with this sub-population.

7 In all three replication sites, the grantees worked with *SSI*'s developer during the grant proposal phase. Upon award, OAH recommended that the developer be involved, and each of the replication sites established consulting agreements for this purpose. The developer was actively involved with each of the grantee sites at the outset and provided the initial training, along with responses to frequently asked questions she received from grantees prior to the availability of the curriculum and implementation materials.

8 In each of the three sites, the video was replaced by one that updated the material and, in some cases, better reflected the racial/ethnic composition of the population served. The developer provided the following guidance for selection of a substitute: brief, include peers, demonstrate correct condom use, and preferably use humor or otherwise be entertaining (correspondence from Lydia Shrier, September 22, 2011). A more detailed description of the videos and adaptations is included in the forthcoming full *SSI* implementation report.

homeless youth. Clinics were located throughout the county in eight cities. The clinics varied in geographic location (urban versus suburban) and the populations served. Each clinic had individual targets for recruitment based on its number of full-time health educators (some health educators were the equivalent of half of an FTE).

Knox County Health Department partnered with two large health agencies (Cherokee Health Systems and Rural Medical Services) to deliver *SSI* in 17 clinics across five counties in eastern Tennessee. The Health Department operates a main office and three satellite offices in Knox County. The partner agencies have offices located in Knox County and the surrounding counties, and most function as regional resources serving residents from across the eastern Tennessee area. Knox County health educators delivered the program in eight of these clinics in Knoxville. Partners delivered the program in community health centers in outlying areas of Knox County and in three adjacent rural counties. Cherokee Health Systems, which oversaw four clinics implementing *SSI*, is a Federally Qualified Community Health Center providing services to rural, poor, and underinsured populations throughout Tennessee, including Knoxville and outlying areas. Rural Medical Services is a Community and Migrant Health Center with five freestanding clinics and one mobile clinic in rural eastern Tennessee countiesPlanned Parenthood of Greater Orlando was the smallest replication, with two clinics. The clinics were located on the west side and the east side of Orlando. The clinics varied in accessibility and by the age and level of risk of the populations each served.

3.5.4 Treatment and Control Conditions

In each of the replication sites, members of the treatment group were offered the initial session of *SSI* and the booster sessions at one-, three-, and six-month intervals, delivered by trained health educators. Members of the control group received the standard of care offered in the clinic or, in the case of PPGO, a choice of either a pregnancy test or an STI test for young women recruited outside the clinic. Both groups could receive non-program services and informational materials offered by the clinic or available in the community. Clinic staff offered members of both the treatment and control groups contraceptive information, but the clinics varied in the level of contraceptive information offered to members of the control group as part of its usual standard of care (Exhibit 3.3).

Exhibit 3.3: Treatment and Control Conditions in the Three Replication Sites

Grantee/Locations	Treatment Group	Control Group
Hennepin County Human Services and Public Health Department 19 clinics in 8 cities in Hennepin County, MN	Individualized sessions with trained health educator; initial session and boosters at 1-, 3-, and 6-month intervals	Standard of care
Knox County Health Department 17 clinics in 5 counties in eastern TN		Standard of care
Planned Parenthood of Greater Orlando 2 clinics in Orlando, FL		Standard services for those recruited in the clinic; pregnancy or STI test for those recruited outside the clinic

3.5.5 Recruitment and Random Assignment

Procedures for the identification and enrollment of young women into the study were similar across replication sites and clinics. In each of the replication sites, potential study participants were identified at the time they came to the clinic for services. Potential study participants, who were seen by clinical staff for scheduled appointments or walk-in (unscheduled) visits, were referred to the health educator. A potential participant might have been a new patient (first time at the clinic or not seen at the clinic for several years) or an established patient (had recently received services at the clinic).

Medical and demographic information was collected by clinic staff for all patients as part of standard clinic procedures. Clinic staff used demographic and other information provided (such as whether the adolescent was currently sexually active and/or pregnant) to screen for eligibility and notify the *SSI* health educators about eligible study participants.

Once an eligible young woman was identified, a health educator made an initial contact to introduce the study and determine whether the young woman was interested. If the young woman was interested, a health educator scheduled her for an in-person enrollment appointment. During this second meeting, a *SSI* health educator described the study and obtained informed consent.[9] All patients—regardless of whether or not they were eligible for study participation, and whether or not they accepted or declined participation in the study—were able to receive the clinic services they requested, according to the standard of care. The structure of intake and enrollment was such that young women generally received the clinical services they sought prior to receiving *SSI* services, although there was some variation in the timing of program receipt.

Young women who consented to the study were then asked to complete the baseline survey. Because intake and random assignment were done individually, on a rolling basis, everyone who provided consent completed a baseline survey. To administer the survey, the health educator logged onto a web-based survey system and then left the respondent in private to complete the survey. Once the baseline survey was completed, the participant was randomly assigned by the health educator to the treatment or control group through a centralized web-based Participant Tracking System (PTS) developed for the study.[10] The random assignment process was designed, managed, and implemented by the Abt study team through the PTS in order to protect the integrity of the random assignment process while allowing health educators to retrieve participants' assignments instantly after completion of the baseline survey. Once the health educator retrieved the results of random assignment, the health educator gave the young woman her gift card for survey completion and informed her of the assignment (treatment or control) and next steps.

Random assignment occurred independently in each of the clinic sites. Individual sample members within clinics were randomly assigned on a rolling basis within randomization blocks based on site and age (younger than 15 versus 15 years or older) and time (3- to 6-month periods). The randomization procedure produced an approximate 2:1 treatment-to-control ratio within the site, age, and time blocks. The random assignment algorithm was programmed by the Abt study team. Program staff members were blind to the algorithm and not able to change the assignment for any individual once it was made. The PTS stored identifying information to ensure that an individual's random assignment status was preserved. In order to minimize crossover from the control group to the treatment group, the PTS was integrated across the clinics within each replication site and checked for duplicate participants across the

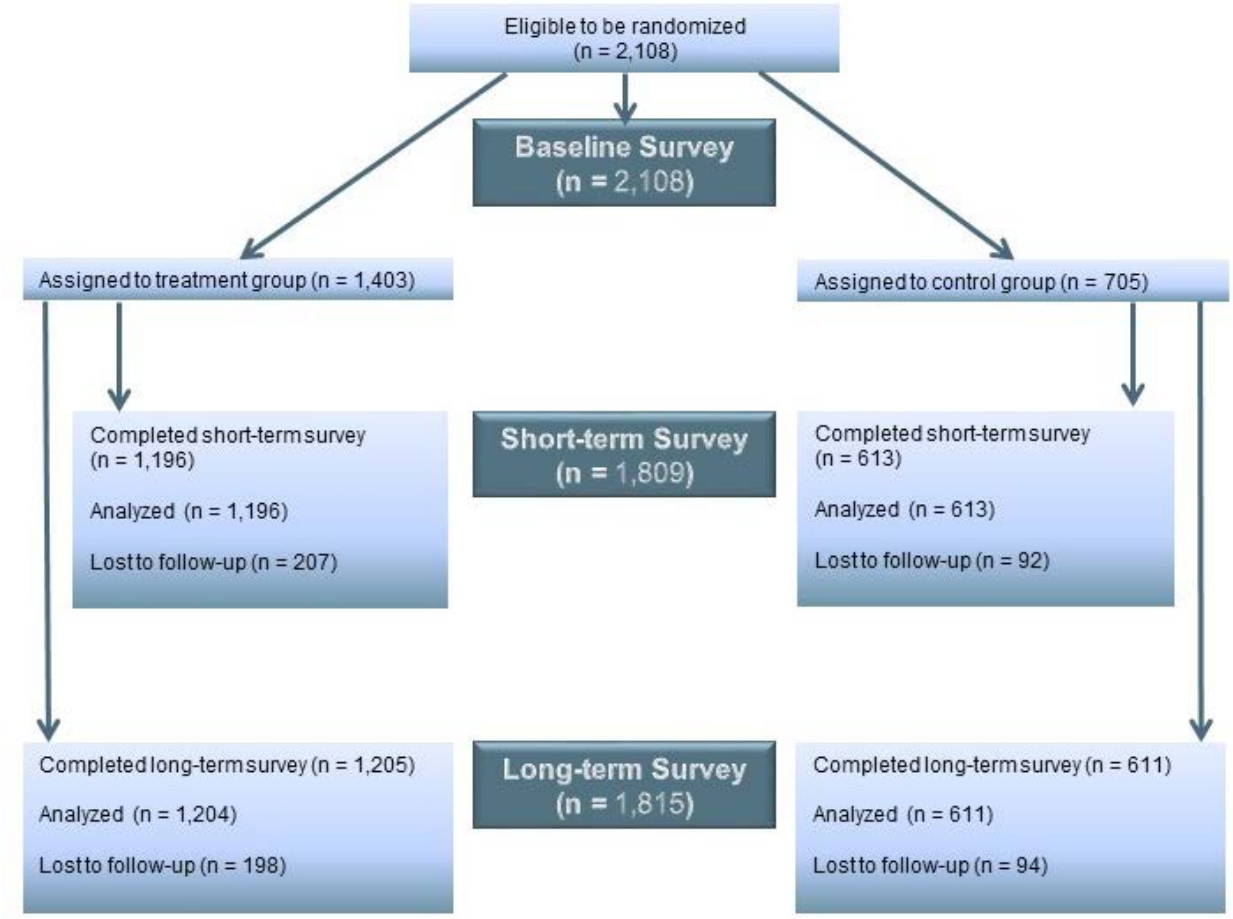
[9] Although participation in the *SSI* program by minor females did not require parental consent, an Institutional Review Board (IRB) waiver was needed to recruit them into the study without parental consent. A small portion of eligible females were accompanied by parents, allowing clinics to seek parental consent for the minor. The study procedures followed the clinic procedures in obtaining parent permission. For minors unaccompanied by a parent, the study obtained a waiver of parent permission from the Abt Associates IRB.

[10] The PTS allowed health educators to conduct random assignment "on the spot" with a fully automated and user-friendly process. The PTS was also used to track participant receipt of booster sessions, monitor fidelity, and notify program intake staff if a potential study member was already enrolled in the study or if a study member in the control group sought *SSI* services.

clinics, so that individuals who had been assigned to the control group at one clinic could not seek out *SSI* services at a different clinic during the study period.

As Exhibit 3.4 shows, across the three replication sites, 2,108 young women were eligible for and consented to the study.[11]

Exhibit 3.4: Study Sample



3.5.6 Data Collection Strategy

A web-based Audio Computer-Assisted Self-Interview (ACASI) system was used to capture and store survey responses, and respondents could choose to take the survey in Spanish or English. At baseline, paper copies of the survey (in Spanish and English) were available as backup in case of computer or Internet failure.

The 30-minute baseline survey was completed individually at each clinic on a computer dedicated to the study. Health educators oversaw the baseline survey and provided gift cards afterward. As Exhibit 3.3 shows, all 2,108 study participants completed a baseline survey.

[11] Data were not collected on youth who declined to participate in the study. Therefore, it is not possible to assess similarities and differences between youth who consented and those who did not.

For the short-term and longer-term follow-up surveys (nine and 18 months after baseline), only the web-based ACASI system was used. For tracking purposes and to invite/remind youth to complete their survey, youth were sent email and text messages before the survey went live and throughout the survey period.[12] Participants were emailed a unique link to the 30-minute follow-up survey, which they completed on their own in any location that was convenient for them, using personal tablets or computers, library computers, or even their smart phones. In some cases, before the survey period closed, field staff contacted participants and encouraged them to complete the survey independently online or helped them to access the survey. Gift cards were mailed to participants after completion.

As Exhibit 3.5 shows, a large majority (85.8 percent) of these young women subsequently completed the short-term survey (nine months after baseline) and 86.1 percent completed the longer-term follow-up survey (18 months after baseline).[13] At both data collection points, there was almost no difference in the response rates of youth in the treatment group versus those in the control group. Response rates varied among the replication sites. Of the three sites, PPGO had the highest response rates (97.3 percent).

Exhibit 3.5: *SSI* Survey Response Rates

	Participants			Completed Short-Term Follow-Up						Completed Longer-Term Follow-Up					
				Total		Treatment		Control		Total		Treatment		Control	
	Total N	Treatment	Control	N	%	N	%	N	%	N	%	N	%	N	%
All Sites	2,108	1,403	705	1,809	85.8	1,196	85.3	613	87.0	1,815	86.1	1,204	85.8	611	86.7
Hennepin County	1,177	785	392	968	82.2	639	81.4	329	83.9	958	81.4	639	81.4	319	81.4
Knox County	491	326	165	413	84.1	275	84.4	138	83.6	429	87.4	281	86.2	148	89.7
Planned Parenthood of Greater Orlando	440	292	148	428	97.3	282	96.6	146	98.7	428	97.3	284	97.3	144	97.3

3.6 Conducting the Analyses

In this section we describe in greater detail the analytic procedure used to address the primary and secondary research questions.

3.6.1 Estimation of Impacts for the Full Sample

We estimated program impacts by comparing the outcomes of treatment and control group members using a regression framework, in which we included baseline covariates to increase statistical precision (i.e., reduce the standard errors) of the impact estimates for a given sample size (Orr, 1999) and reduce attrition bias from missing data (see Puma, Olsen, Bell, & Price, 2009). For each outcome measure, the model produces an estimate of the average treatment impact of *SSI* across the three replication sites.

Individual sample members were randomly assigned within randomization blocks based on site, clinic, age (less than 15 years versus 15 years or older), and time (three- to six-month periods) in an approximate 2:1 treatment-to-control ratio. The model includes indicator variables for each randomization block to

[12] Participants were allowed a three-month window to complete the follow-up survey.

[13] Individuals were excluded from the analysis if more than 75 percent of survey items were missing data. One individual in the treatment group was excluded from the longer-term follow-up for this reason.

compare treatment and control group members within site, clinic, age, and time and to account for the unequal assignment ratio. The estimated impact is therefore a precision weighted average of the estimated treatment effects within randomization blocks. For each outcome, we estimate a model that reflects this design and has the basic structure of Equation 1.[14]

Eq (1) $$Y_i = \beta_0 + \beta_1 T_i + \sum_{k=1}^{K} \lambda_k X_{ki} + \sum_{m=1}^{M} \gamma_{0m} D_m + \varepsilon_i$$

In this model:[15]

Y_i is the outcome of interest (e.g., sexual intercourse without birth control) for the i^{th} individual in the m^{th} randomization block.

T_i is an indicator variable equal to 1 if individual *i* was assigned to the treatment group and 0 otherwise.

X_{ki} is the k^{th} baseline covariate; these include baseline age, race/ethnicity (Black, White, Hispanic (omitted), other), risk behaviors (smoking, alcohol use, marijuana use), baseline sexual activity (ever sexually active), baseline pregnancy risk knowledge, baseline STI risk knowledge, baseline intentions to have oral sex, baseline intentions to have sexual intercourse, and the baseline measure of the outcome when available. For the longer-term follow-up, ever pregnant at baseline was also included in the model as a baseline covariate.

D_m is the indicator variable representing the m^{th} randomization block.

ε_i is the usual random error term.

In this model, β_1 represents the average pooled impact of the program on the outcome. The *p*-values reported for impact estimates are two-tailed to account for the possibility that the intervention might adversely affect one or more of the outcomes. The coefficients on the covariates, λ_k, reflect the relationship between the outcome measure and each of the covariates while controlling for others. It is important to note that this model specification treats randomization blocks (and thus sites) and the treatment effects as fixed as opposed to random, which is consistent with how the replication sites were chosen and how the results of the study will be interpreted.[16]

Equation 1 estimates the impact of assignment to *SSI*. The crucial difference between the treatment and control groups is ***access*** to *SSI* services: Individuals in the treatment group had access to program services and potentially similar information in the clinics, as well as access to other services in the community;

---

[14] Because random assignment occurred at the individual level (not the clinic level) within randomization blocks, we estimated a one-level fixed-effects model that included a series of indicator variables representing each of the randomization blocks defined by site, clinic, age, and time (Bloom, 2006, p. 13).

[15] The analyses presented in this report used linear probability models for binary outcomes for interpretability. A set of robustness analyses were conducted using logistic regression models and using linear models with heteroskedasticity robust standard errors for binary outcomes (Constantine et al., 2009; Gleason, Clark, Tuttle, & Dwyer, 2010). There were no substantive differences in the inferences that results from any of the three modeling approaches.

[16] Because replication sites were selected as a purposive sample, not randomly selected from a larger population of sites, we do not consider a random treatment effects model to be appropriate for drawing inferences from this sample (Schochet, 2008b, p. 70).

control group members had access to the standard services available in the clinics and other services in the community. The analyses estimate the impact of having the opportunity to participate in the intervention, not the average impact on program group members who actually participate in the intervention. In *SSI*, where there was a very high rate of participation in intervention services by members of the treatment group, this estimate will be very close to the impact on the members who actually participated in services.

3.6.2 Correcting for the Number of Comparisons Needed to Answer Key Questions

The confirmatory analyses estimate impacts on the key outcomes for the full sample, using data pooled across the three replication sites. Prioritizing the analyses limits the number of hypothesis tests we conduct to draw causal conclusions. Typically, we use a *p*-value criterion of .05 to determine whether an impact estimate is statistically significant and unlikely to be a chance finding. Limiting the number of hypothesis tests we conducted helps mitigate the risk of incorrectly concluding that *SSI* had an impact.

However, we also applied a correction for multiple comparisons within the key outcome domains that had more than one outcome measure. Within each of the two measurement domains identified in Exhibit 3.2 that have more than one outcome measure (*recent sexual behavior at the short-term follow-up; recent sexual behavior at the longer-term follow-up*), we applied a correction described by Benjamini & Hochberg (1995) that adjusts the criterion used for determining statistical significance to account for multiple tests. In this case, the correction means that within an outcome domain that included two key outcome measures, both of the tests would be deemed significant if both have *p*-values below .05; if only one has a *p*-value below .05, it would be deemed significant only if its *p*-value is below .025. For the *consequences of sexual risk behavior* outcome domain, there was only one outcome measure, so no multiple comparisons correction was applied. In this domain, we applied the traditional criterion for statistical significance of $p < .05$.

For exploratory analyses (i.e., all non-confirmatory analyses), we applied no adjustments to the criterion for statistical significance.[17] For each exploratory test, we applied the traditional criterion for statistical significance of $p < .05$. As noted previously, exploratory analyses should not be used to make causal conclusions about the effectiveness of *SSI*. The results from exploratory analyses are reported separately from the results of confirmatory analyses, and readers should interpret those results with caution keeping in mind that with a large number of tests conducted, the likelihood of obtaining statistically significant results by chance is high. Even if there were no true impact of the intervention on participants, we would expect that five percent of the tests would be significant by chance alone.

3.6.3 Site-Level Analyses

We also estimated effects for each site separately and tested for differences in effects between the three sites by including treatment-by-site interaction terms in Equation 1 above (see Section 3.6.1) and testing

[17] The decision not to apply an adjustment for multiple comparisons to the results of the exploratory analyses aligns with standards of good practice (see Schochet, 2008a) and was made after weighing the risks and benefits. The risk of not applying adjustments for multiple comparisons in the exploratory analyses is the likelihood of spurious findings. Conversely, if we were to apply multiple comparisons adjustments to the exploratory findings, the adjustments would be very conservative and practically no results would be flagged as significant. The benefit of reporting unadjusted results from the exploratory analyses is that unadjusted test results help to identify potentially important findings that may, in turn, help interpret the findings from the confirmatory analyses; doing so also may suggest promising avenues for future research.

for the joint significance of the interaction terms. When statistically significant differences in impact are found between sites for one or more outcomes, we discuss these differences. The purpose of testing for differences between sites before discussing site-level results in the main text is to guard against overinterpretation of spurious findings, some of which would be expected by chance in such a large group of outcomes. We discuss site-specific effects only when differences in effects between sites are found, because it is only credible to report an effect in one site—but not in another—if there is a significant difference between the sites. The site-specific results in Appendix B are not adjusted for multiple comparisons, and any significant findings reported there should be interpreted with caution.

3.6.4 Subgroup Analyses

In addition to the overall and site-level effects, we estimated effects for key subgroups of participants—based on age (younger than 15, age 15 or older), race/ethnicity (Hispanic, Black, White, Other), and baseline sexual experience (never sexually active at baseline / ever sexually active at baseline)—and tested for differences between subgroups, to better understand what works for whom. We implemented subgroup analyses by including subgroup indicators and treatment-by-subgroup interaction terms in the model (i.e., Equation 1 above in section 3.6.1) and testing for the significance of the interaction term.

To reduce the potential for overinterpretation of results among the large number of subgroup estimates, we present impact estimates for individual subgroups in Appendix C when there is a statistically significant difference between subgroups; for example, the impact would be presented for the subgroup of younger participants only if there were a statistically significant difference in impacts between younger and older participants.

3.6.5 Approach to Handling Missing Data

We used case deletion for the few instances of missing outcome data (Puma et al., 2009). Dummy-variable adjustment was used in regression models to account for missing covariates. In the dummy variable adjustment method, missing covariate values were set to a constant and indicators (or dummy variables) for such values were added to the impact analysis model (Puma et al., 2009).

4. Implementation Findings and Baseline Characteristics

Before presenting the impact results of the study, we first consider some important contextual factors that might affect the interpretation of the findings. How well a program model is implemented, as well as the characteristics of the population served, can strongly influence the extent to which the program is able to meet its goals.

Implementation of the *SSI* program was guided by fidelity requirements established by OAH at the outset of the grant award. The guidelines allow an assessment of the extent to which the program was implemented with fidelity and to highlight areas where there were differences in implementation across the replication sites. In this chapter, we expand on our conclusion that the intervention was implemented with fidelity across replication sites and describe the study sample at baseline (i.e., when students were enrolled in the study).[18]

Key Findings

Across the three replication sites,

- *SSI* was implemented as intended, and participants received a majority of the intervention.
- After some initial challenges, each of the replication sites successfully enrolled and served the intended population of young women.
- Across sites there was variation in demographic characteristics of participants, engagement in risk behaviors at baseline, and intentions to engage in risky behavior.

4.1 Sample Characteristics

In this section, we present baseline characteristics of the analytic samples pooled across all three sites as well as for each individual site. We then describe the comparability of the treatment and control groups at baseline.

4.1.1 Analytic Samples

Baseline characteristics of the analytic sample for *SSI* overall and for each replication site are presented in Exhibit 4.1.[19]

Age. At baseline, the young women in the study sample were, on average, 17.2 years old.

Race/Ethnicity. More than one-third of participants were non-Hispanic Black, almost one-third were White, and the remaining third were nearly equally divided between Hispanic and Other (which includes Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, multiracial, and undisclosed race).

Family Structure. Across all three replication sites, more than 75 percent of youth lived with one or both biological parents. Overall, fewer than half said that they felt very close to and cared for by their mothers, and fewer than a third reported they felt close to and cared for by their fathers.

Risk Behaviors. Across all three sites, more than three-quarters had ever drank alcohol, more than two-thirds had ever used marijuana, and slightly more than half had ever smoked cigarettes.

---

[18] A more detailed description and analysis of implementation is provided in the forthcoming full implementation report.

[19] Because of very low attrition, the baseline characteristics of the short-term analytic sample differ little if at all from the characteristics of the longer-term analytic sample shown in Exhibit 4.1. For interested readers, the baseline characteristics of the short-term analytic sample are shown in Appendix Table D.2

Sexual Activity/Risk Behavior/Consequences. Given the eligibility criteria for the program, it is not surprising that at the time of study enrollment most of the participants were sexually active and intended to be sexually active in the next 12 months. More than 90 percent of the sample had ever been sexually active (engaged in sexual intercourse, oral sex, and/or anal sex), and 83 percent had been sexually active in the 90 days before the study began. In the 90 days before the study began, 79 percent had engaged in sexual intercourse, 66 percent had engaged in oral sex, and 11 percent had engaged in anal sex.

During that same period, a large proportion of the sample had engaged in unprotected sexual activity. More than 60 percent of all participants had engaged in sexual intercourse without a condom or oral sex without a condom, and about nine percent of all participants had engaged in anal sex without a condom. At baseline, nearly one-fifth of the sample had ever been pregnant. More than one-tenth of the sample reported being diagnosed with an STI in the past year.

Knowledge/Attitudes/Intentions. Participants in all three replication sites were knowledgeable about pregnancy risk factors and STI risk factors, with average knowledge scores around 70 out of 100. More than 80 percent of participants in all three replications sites intended to have sexual intercourse in the next 12 months and to use condoms if they did. Overall, more than 90 percent of participants intended to use birth control if they had sexual intercourse in the next 12 months.

Exhibit 4.1: Baseline Characteristics of the Longer-Term Analytic Sample by Site

Measure	Hennepin County	Knox County	Planned Parenthood of Greater Orlando	*SSI* Overall	*p*-Value for the Test of Differences Across Sites[a]
Demographic characteristics					
Age (years)					
Mean	16.96	17.14	17.58	17.15	.000***
Race/ethnicity[b]					
Hispanic	17.12	9.09	27.57	17.69	.000***
Non-Hispanic Black	36.22	24.71	45.79	35.76	.000***
White	25.89	60.14	21.50	32.95	.000***
Other	20.77	6.06	5.14	13.61	.000***
Family structure and relationships					
Lives with biological parent/s	81.49	75.18	75.76	78.65	.008**
Feels very close to and cared for by mother	41.40	55.19	44.94	45.51	.000***
Feels very close to and cared for by father	23.91	35.68	27.15	27.48	.000***
Risk behaviors					
Ever smoked cigarettes	54.39	59.06	42.29	52.63	.000***
Ever drank alcohol	80.19	78.12	83.41	80.46	.142
Ever used marijuana	72.98	60.47	62.53	67.57	.000***
Sexual activity					
Ever sexually active[c]	95.06	90.19	94.37	93.74	.002**
Recently sexually active (in the last 90 days) [c]	86.11	79.58	80.52	83.24	.003**
Sexual intercourse in the last 90 days	82.30	75.47	75.12	78.98	.001**
Oral sex in the last 90 days	66.42	64.55	65.26	65.70	.778
Anal sex in the last 90 days	11.70	na	9.39	10.98	.206
Sexual risk behavior [c]					
Sexual intercourse without a condom in the last 90 days	65.23	56.54	54.93	60.73	.000***
Oral sex without a condom in the last 90 days	62.72	62.21	59.62	61.87	.543
Anal sex without a condom in the last 90 days	9.06		8.22	8.80	.609

Measure	Hennepin County	Knox County	Planned Parenthood of Greater Orlando	*SSI* Overall	*p*-Value for the Test of Differences Across Sites[a]
Sexual intercourse without birth control in the last 90 days	32.24	28.50	36.15	32.28	.058
Sexual intercourse with more than one partner (lifetime)	66.91	63.68	67.61	66.31	.409
Sexual intercourse with more than five partners (lifetime)	24.36	20.99	23.88	23.45	.386
Consequences of sexual risk behavior					
Ever pregnant (lifetime)	18.14	18.59	18.54	18.34	.973
Diagnosed with STI in the last 12 months	16.95	10.33	6.57	12.93	.000***
Knowledge, attitudes and intentions					
Knowledge of pregnancy risk[d]	68.29	69.74	70.84	69.23	.473
Knowledge of STI risk[d]	67.97	66.26	70.37	68.14	.086
Attitudes toward protection[e]	3.25	3.28	3.22	3.25	.042*
Intentions to have oral sex[f]	60.02	58.55	62.91	60.35	.409
Intentions to have sexual intercourse[f]	85.97	80.09	83.88	84.09	.022*
Intentions to use a condom if they were to have sexual intercourse[f]	83.75	82.24	86.21	83.98	.275
Intentions to use birth control if they were to have sexual intercourse[f]	93.09	95.32	86.68	92.10	.000***

Source: Baseline survey administered prior to randomization.
Note: Data in this table are based on 1,385–1,809 longer-term survey respondents who provided valid survey responses to relevant items on the baseline survey. Values shown are percentages unless otherwise indicated. The items that compose measures of attitudes toward risky sexual behavior, motivation to delay childbearing, refusal skills, and condom negotiation skills were not asked at baseline. na is not asked.
[a] Test results from an analysis of variance testing the null hypothesis that the means of the variable indicated in the row are equivalent among the three sites.
[b] *Other* is defined as Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, multiracial, or undisclosed race.
[c] *Sexual activity* is defined differently across grantees. In Hennepin County and PPGO, sexual activity refers to sexual intercourse, oral sex, and anal sex. Youth were not asked about anal sex in Knox County.
[d] Scores represent the average percentage of items answered correctly.
[e] Scale score averages responses ranging from 1 to 4. Higher scores indicate more positive attitudes.
[f] Intention to engage in the behavior in the next 12 months. Dichotomous variables, reported as percentage of respondents who responded affirmatively.
* $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed tests).

Differences Among Sites. The site-specific profiles at baseline of youth participating in the study differ from one another in several ways. Hennepin County was more ethnically/racially diverse than the other replication sites: One-fifth of participants were of Other race, more than a third were Black, and slightly more than 25 percent were White. By contrast, Knox County participants were predominantly White, with one-quarter Black, less than 10 percent Hispanic, and less than 10 percent Other. Almost half of the participants in PPGO were Black, more than 25 percent were Hispanic, and about one-fifth were White.

In general, young women in Knox County appeared at lower risk relative to their counterparts in Hennepin County and PPGO on several indicators at baseline; specifically, in their use of marijuana, attitudes toward protection, intentions to have sexual intercourse, intentions to use birth control, sexual initiation (ever sexually active), and feelings of closeness with their parents. Rates of sexual activity, sexual intercourse without a condom in the last 90 days and diagnosis with an STI in the last year were highest in Hennepin County. Young women in PPGO reported lower levels of exposure to information about contraceptives (i.e., birth control methods, where to obtain birth control, and how to talk with partner about sex and birth control) at baseline than in the other two replication sites.

4.1.2 Comparability of the Treatment and Control Groups at Baseline

Although the characteristics of study participants differed significantly across the three replication sites (reflecting the differences in youth populations in those sites), there were almost no significant differences between those assigned to the treatment group and those assigned to the control group (Appendix Exhibits D.1 and D.2).

Baseline treatment-control differences were estimated for both the short-term and longer-term analytic samples using a series of models with the same structural components as the impact model in Equation 1 (i.e., the same randomization block indicators and treatment group indicator), but in each model one baseline characteristic (from among those in Exhibit 4.1) served as the dependent variable, and the other covariates used in the impact model (e.g., race/ethnicity, age, ever sexually active) were omitted. In this approach, the coefficient for the treatment indicator is the treatment-control difference on the pre-test measure.

For the longer-term analytic sample, at baseline, there were two significant differences between the two groups (see Appendix Table D.1). Fewer young women in the treatment group reported ever having drunk alcohol than did young women in the control group, and a greater number of young women in the treatment group reported ever getting pregnant. There were no significant differences between the two groups on any of the measures at the short-term analytic sample (see Appendix Table D.2). Variables for which there were differences were subsequently included in the impact models as covariates.

4.2 Program Implementation

As we noted at the beginning of this report, a forthcoming report will provide a detailed account of the implementation of *SSI* in the three replication sites. That implementation report serves two important purposes: (1) to help explain the findings of the impact study and (2) to offer lessons learned to help those planning to use the *SSI* program in the future. In this section, we provide a summary of the implementation findings that are directly relevant to the impact findings reported in the next chapters.

SSI was generally well implemented across the three replications. The three grantees hired staff with appropriate background experience and skills to deliver the program; all staff received training approved by the developer; the program was implemented with fidelity to its core elements and without modifications that threatened those core elements; and attendance was generally strong.

4.2.1 Staff Hiring and Training

The three grantees were consistent in the types of experience and skills they sought when hiring health educators (or identifying one or more from current clinic staff). Experience working with adolescents and in sexual health and comfort in addressing adolescent sexual health issues were considered important.[20] All of the replication sites stressed the importance of being comfortable with the program content and approach. In Hennepin County, program leadership sought individuals who were committed to the rationale for motivational interviewing and who understood that the health educator role was more about listening and eliciting conversation than about teaching. Project staff from each of the three grantees (supervisors and selected health educators) attended a two-day training led by the program developer. Attendees were then responsible for training other health educators.

Most health educators did not have formal training in motivational interviewing; in all three sites, program managers developed additional training specifically to supplement the intervention materials on

[20] Education or training in sexual health was not a requirement in PPGO.

motivational interviewing. PPGO developed an extensive two-week training, with motivational interviewing as a primary focus. Hennepin County provided in-service training on a wide range of topics, including working with youth, dealing with sexual assault, and ethics and boundaries of youth workers. To the extent feasible, health educators in each of the replication sites attended periodic training sessions offered by OAH and were encouraged to seek additional training.

4.2.2 Implementing the Program with Fidelity

As part of the TPP Program, OAH stipulated that grantees maintain fidelity to the core components of their chosen program model, and it provided guidance on making minor adaptations (all of which had to be approved by OAH before they could be implemented). There was an accompanying requirement that grantees develop a plan to monitor fidelity of implementation and continued adherence to the core program model.

For *SSI*, fidelity monitoring checklists were provided by the developer to help grantees collect this information. Health educators were required to complete a fidelity log for each session delivered. Data from the fidelity logs were aggregated and used by program supervisors to identify areas where improvement was needed. Given the personalized and private nature of the intervention, OAH waived the requirement for observations. Each of the replication sites developed processes for monitoring the performance of health educators—in at least one site this was through observations of mock sessions using youth actors or, in some cases, other program staff. Aggregate data on fidelity were delivered to OAH every six months and summarized to provide a basis for subsequent discussions between program officers and the grantees. All of these activities were intended to guide implementation and ensure not just fidelity but also a degree of uniformity across sites replicating the same program model.

Each of the replication sites successfully delivered the intervention to youth with fidelity to the program model. Nevertheless, grantees discovered they needed to develop strategies to address implementation challenges. Each of the grantees struggled to some extent with retention. Young women did not always attend all of the booster sessions, and each of the replication sites developed strategies to address what it perceived to be the reasons why. PPGO identified transportation as a substantial barrier to participation, so it hired a transportation company to transport young women to and from sessions. In addition, PPGO received approval from OAH to offer booster sessions remotely via video conference or smart phone video chat (e.g., Skype, FaceTime). The other two replication sites also received approval for remote video for the booster sessions, but they were less successful in implementing this adaptation. Knox County also extended clinic hours to accommodate young women's schedules.

4.2.3 Participant Attendance and Engagement

Grantees were required to collect and report participant attendance (by session). Attendance rates differed slightly by replication site. Roughly 60 percent of participants in Knox County and Hennepin County attended 75 percent or more of the sessions. The median number of sessions for both of these sites was 2.7 (out of 4). The numbers were slightly higher in PPGO, where 67 percent attended 75 percent or more of the sessions, and the median number of sessions was 3.0 (out of 4).

The PTS (see Section 3.5.5) was developed both to meet the needs of the study and to allow grantees to collect program monitoring data, including attendance and fidelity. Health educators had individual login credentials and entered the fidelity and participation information for the sessions they delivered directly into the system. Supervisors used the PTS to generate reports on attendance and fidelity.

5. Program Impacts on Youth Sexual Activity, Sexual Risk Behavior, and Consequences of Sexual Risk Behavior

In this section of the report, we present findings for both the short-term and longer-term follow-up surveys on the behavioral outcomes of interest. The findings presented here reflect our analytic strategy of first conducting confirmatory analyses by examining a key set of outcomes for the pooled sample to produce results that are more conclusive about the impacts of *SSI* rather than suggestive.

We begin this chapter with a discussion of the confirmatory analyses, followed by a presentation of program effects on other related sexual risk behaviors and consequences for the full sample. Findings for site-level effects and specific subgroups of interest are described in Chapter 7.

5.1 Confirmatory Analyses of Impacts on Key Behavioral Outcomes

The pre-specified confirmatory analyses test the impacts of *SSI* on the following key outcomes for the full sample: *currently sexually active* and *sexual intercourse without birth control in the short-term* (at 9 months); *currently sexually active* and *sexual intercourse without birth control at the longer-term* (at 18 months); and *pregnancy* (between the baseline and 18-month follow-up survey). In order to minimize the concern that our confirmatory analysis would miss a behavioral impact that occurred early in the follow-up period but nonetheless affected pregnancy, we treat recent sexual behavior at the short-term follow-up as distinct from recent sexual behavior at the longer-term follow-up.

After nine months, *SSI* significantly reduced reported sexual intercourse without birth control. However, at the longer-term follow-up (after 18 months), this difference is no longer statistically significant.

On average, at the short-term *SSI* participants were 5.8 percentage points (21 percent) less likely to report engaging in sexual intercourse without birth control than the control group were. At the longer-term follow-up, this gap is reduced by half: *SSI* participants were 2.9 percentage points (11 percent) less likely to report engaging in sexual intercourse without birth control. There were no impacts on current sexual activity at either the short-term or at the longer-term follow-up. After 18 months, slightly more than three-quarters of participants in both groups reported engaging in sexual activity within the last 90 days

The final confirmatory analysis tested the program's impact on pregnancy in the period between baseline and the longer-term follow-up. After 18 months, 16 percent of youth in the treatment group and 19.4 percent of youth in the control group reported getting pregnant since the baseline. Though not quite reaching the established criterion for statistical significance of $p < .05$, the finding does favor the treatment group.

Key Behavioral Impact Findings

- Confirmatory analyses revealed significant impacts of *SSI* on sexual intercourse without birth control (9 months after baseline).
- The program had an impact on pregnancy that, while not statistically significant, was promising.
- The program had no impacts on recent sexual activity at either the short-term or longer-term follow-up.

Exploratory analyses revealed no significant overall effects of *SSI* on:

- Recent sexual intercourse, recent oral sex, or recent anal sex.
- Recent sexual intercourse without a condom; recent oral sex without a condom, recent anal sex without a condom.
- Recent diagnosis of a sexually transmitted infection.

Exhibit 5.1: Short-Term and Longer-Term Impacts on Confirmatory Behavioral Outcomes

Outcome	Short-Term Impacts				Longer-Term Impacts			
	Adjusted Treatment Mean[a]	Unadjusted Control Mean	Treatment Effect[b]	*p*-Value	Adjusted Treatment Mean[a]	Unadjusted Control Mean	Treatment Effect[b]	*p*-Value
Sexual Behavior								
Sexual activity (percentage responding affirmatively)								
Currently sexually active (in the last 90 days)[c]	74.84	74.96	−0.11	.954	75.12	76.11	−0.99	.624
Sexual risk behavior (percentage responding affirmatively)								
Sexual intercourse without birth control (in the last 90 days)	22.05	27.82	−5.78**[d]	.005	23.84	26.69	−2.85	.179
Consequences of sexual risk behavior (percentage responding affirmatively)								
Pregnant since baseline					16.00	19.41	−3.41	.070[e]

Source: Follow-up surveys administered nine and 18 months after baseline.
Note: Short-term results are based on 1,801 respondents who provided valid survey responses to relevant items. Longer-term results are based on 1,806–1,807 respondents who provided valid survey responses to relevant items except for the item *pregnancy since baseline* (*n* = 1,700). Outcomes reported in this table are binary, and we report impacts as percentage point differences between the treatment and control group means.
[a] The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression-adjusted impact estimate (treatment effect).
[b] The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. The treatment effect is expressed as a difference in percentage points. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.
[c] *Sexual activity* is defined differently across grantees. In Hennepin County and Planned Parenthood of Greater Orlando, sexual activity refers to sexual intercourse, oral sex, and anal sex. Youth were not asked about anal sex in Knox County.
[d] Indicates statistical significance after application of Benjamini-Hochberg (1995) correction for two tests within this outcome domain. The criterion for statistical significance is $p < .05$ if both tests have *p*-values less than .05, and is .025 if only one of the two tests has a *p*-value less than .05.
[e] Criterion for statistical significance is $p < .05$.

5.2 Exploratory Analyses of Impacts on Additional Behavioral Outcomes

We also conducted a series of exploratory analyses that, though only suggestive of evidence of program effectiveness, are supported by theory (the program logic model), are supported by the experimental study design, and were specified in advance of the analysis. These analyses tested program effects on other sexual behaviors, sexual risk behaviors, and consequences for the full sample.

***SSI* had no statistically significant effects on the prevalence of sexual behaviors at nine or 18 months after study enrollment.**

As shown in Exhibit 5.2, slightly less than three-quarters of young women in both the treatment and control groups reported that they had sexual intercourse in the last 90 days at the 18-month follow-up; slightly more than 60 percent reported having recently engaged in oral sex; and nearly 10 percent reported engaging in anal sex.

We also found no evidence of program effects on rates of sexual risk behaviors, including sexual intercourse, oral sex, or anal sex without a condom. At the longer-term follow-up, 55 percent of youth in the treatment group and 59 percent of youth in the control group reported having sexual intercourse without a condom in the past 90 days; slightly less than 60 percent of young women in both groups reported engaging in recent oral sex without a condom; and a few young women (less than 10 percent) in both groups reported engaging in anal sex without a condom.

At the time of the longer-term follow-up, *SSI* had no effect on STI diagnoses. Approximately 10 percent of young women in the treatment and control groups reported being diagnosed with an STI in the prior 12 months.

Exhibit 5.2: Additional Short-Term and Longer-Term Effects on Sexual Activity, Sexual Risk Behavior, and Consequences

Outcome	Short-Term Impacts				Longer-Term Impacts			
	Adjusted Treatment Mean[a]	Unadjusted Control Mean	Treatment Effect[b]	*p*-Value	Adjusted Treatment Mean[a]	Unadjusted Control Mean	Treatment Effect[b]	*p*-Value
Sexual Behavior								
Sexual activity (percentage responding affirmatively)								
Sexual intercourse in the last 90 days	71.29	72.18	−0.89	.661	71.84	72.49	−0.64	.755
Oral sex in the last 90 days	59.32	60.39	−1.07	.626	60.60	61.29	−0.68	.759
Anal sex in the last 90 days[c]	9.13	6.13	2.99	.051	9.13	10.00	−0.87	.597
Sexual risk behavior (percentage responding affirmatively)								
Sexual intercourse without a condom (in the last 90 days)	53.66	57.45	−3.79	.087	55.45	58.98	−3.52	.128
Oral sex without a condom (in the last 90 days)	54.32	56.63	−2.31	.299	56.23	57.66	−1.43	.527
Anal sex without a condom (in the last 90 days)[c]	7.32	4.65	2.67	.056	6.81	8.48	−1.67	.260
Sexual intercourse with more than one partner (lifetime)	70.07	71.82	−1.75	.332	74.30	73.67	0.63	.741
Sexual intercourse with more than five partners (lifetime)	26.35	28.86	−2.51	.163	32.66	31.33	1.33	.503
Consequences of sexual risk behavior (percentage responding affirmatively)								
Diagnosed with STI in the last 12 months	n/a	n/a	n/a	n/a	9.65	11.02	−1.37	.354

Source: Follow−up surveys administered nine and 18 months after baseline.
Note: Short−term results are based on 1,801 respondents who provided valid survey responses to relevant items, except for the items measuring number of partners (*n* = 1,735) and anal sex (*n* = 1,389). Longer-term results are based on 1,806–1,808 respondents who provided valid survey responses to relevant items except for the items measuring anal sex (*n* = 1,379), number of partners (*n* = 1,788), and pregnancy (*n* = 1,700).
[a] The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression-adjusted impact estimate (treatment effect).
[b] The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. The treatment effect is expressed as a difference in percentage points. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.
[c] *Sexual activity* is defined differently across grantees. In Hennepin County and Planned Parenthood of Greater Orlando, sexual activity refers to sexual intercourse, oral sex, and anal sex. Youth were not asked about anal sex in Knox County.

6. Exploratory Analyses of Program Effects on Non-Behavioral Intermediate Outcomes

The *SSI* program theory of change (see logic model in Exhibit 2.3) specifies a set of intermediate outcomes that the model predicts will influence behavior. If the theory underlying the logic model is correct, we would expect to see direct effects on behavior in the short-term, but we would also expect positive effects on these non-behavioral intermediate outcomes in the short term, and that those effects would be sustained over time such that young women change their behavior in ways that ultimately protect them from the potential consequences of sexual risk behavior (e.g., from STIs and early pregnancy).

Accordingly, the study is designed to determine whether *SSI* also affects those non-behavioral outcomes Specifically, when delivered with fidelity, the program is intended to affect young womens' knowledge and understanding of reproductive health and avoidance of sexual risk, attitudes toward using protection, motivation to delay childbearing, intentions to engage in sexual activity and use protection, and skills needed to avoid sexual risk.

Key Findings on Non-Behavioral Outcomes

- *SSI* improved refusal skills after nine months and after 18 months; after 18 months *SSI* also improved condom negotiation skills.
- *SSI* increased intentions to use protection after nine months, but not after 18 months.
- *SSI* improved attitudes towards using protection after nine months and improved attitudes towards extremely risky sexual behaviors after 18 months.
- *SSI* had no effects on knowledge or motivation to delay childbearing

The earlier short-term report[21] presented detailed findings on the impact of *SSI* (nine months after the study began) on a range of non-behavioral intermediate outcomes. In the sections below, we summarize those findings and include findings for the longer-term follow-up.

6.1 Knowledge of Pregnancy and STI Risk

At both the short-term (9 months after baseline) and longer-term (18 months after baseline) follow-up, *SSI* had no statistically significant effects on knowledge of pregnancy risk or knowledge of STI risk.

SSI had no effect on measures of knowledge—either for pregnancy risk or STI risk, at either the short- or longer-term follow-up. In general, participants were well informed about methods of preventing pregnancy and had general knowledge of STI facts, transmission, and prevention. At both time points, study participants in both groups (treatment and control) correctly answered 75 percent or more of the items on the two composite measures of risk (Exhibit 6.1).

[21] Interested readers can find more detailed information at https://aspe.hhs.gov/pdf-report/teen-pregnancy-prevention-replication-study-short-term-impacts-safer-sex-intervention.

Exhibit 6.1: Short-Term and Longer-Term Effects of SSI on Knowledge

Outcome	Adjusted Treatment Mean[a]	Unadjusted Control Mean	Treatment Effect[b]	*p*-Value
Short-Term Follow-Up				
Knowledge of pregnancy risk[c]	78.53	78.26	0.27	.817
Knowledge of STI risk[c]	75.91	74.80	1.11	.183
Longer-Term Follow-Up				
Knowledge of pregnancy risk[c]	77.47	78.03	−0.56	.651
Knowledge of STI risk[c]	75.40	75.75	−0.35	.691

Source: Follow-up surveys administered 9 months and 18 months after baseline.
Note: Results in this table are based on 1,809 respondents (short-term survey) and 1815 respondents (longer-term survey) who provided valid survey responses to relevant items.
[a] The treatment group mean is regression adjusted, calculated as the sum of the control group mean and the regression-adjusted impact estimate (treatment effect).
[b] The treatment effect was estimated in a multi-level model that controls for randomization blocks and other covariates. The treatment effect is expressed in percentage points. Due to rounding, reported treatment effects may differ from differences between reported means for the treatment and control groups.
[c] Scores represent the average percentage of items answered correctly.

6.2 Attitudes

***SSI* had a small but statistically significant effect on the composite measure of participants' attitudes toward using protection (birth control and condoms) nine months after baseline.** Although both groups expressed positive attitudes toward using birth control or condoms, on average, the treatment group had slightly more positive attitudes than did the control group. The treatment effect was the same at both time points, but after 18 months this difference was no longer statistically significant.

The program had no statistically significant effects on youth attitudes toward risky behavior after nine months. Almost all young women in both groups rejected the view that risky behaviors were acceptable. However, after 18 months, participants in the treatment group reported even less support for risky behavior, a difference that was statistically significant (Exhibit 6.2).

Exhibit 6.2: Short-Term and Longer-Term Effects of SSI on Attitudes

Outcome	Adjusted Treatment Mean[a]	Unadjusted Control Mean	Treatment Effect[b]	SES[c]	*p*-Value
Short-Term Follow-Up					
Attitudes toward protection[d]	3.36	3.32	0.03*	0.09	.050
Attitudes toward risky behavior[e]	4.12	5.42	−1.30		.061
Longer-Term Follow-Up					
Attitudes toward protection[d]	3.32	3.29	0.03	0.07	.130
Attitudes toward risky behavior[e]	4.99	6.63	−1.64*		.028

Source: Follow-up surveys administered 9 months and 18 months after baseline.
Note: Results in this table are based on 1,802–1,809 respondents (short-term survey) and 1,810-1,815 respondents (longer-term survey) who provided valid survey responses to relevant items.
[a] The treatment group mean is regression adjusted, calculated as the sum of the control group mean and the regression-adjusted impact estimate (treatment effect).
[b] The treatment effect was estimated in a multi-level model that controls for randomization blocks and other covariates. The treatment effect is expressed in percentage points. Due to rounding, reported treatment effects may differ from differences between reported means for the treatment and control groups.
[c] The SES is the standardized effect size of the difference. For outcomes that are not reported as percentages, the SES is the treatment effect divided by the pooled standard deviation of the treatment and control groups. n/a is not applicable.
[d] Scale score averages responses ranging from 1 to 4. Higher scores indicate more positive attitudes.
[e] Score represents the average percentage of items agreed with (ranging from 0 to 100). Higher values indicate more support for risky sexual behavior.*

p<.05, ** *p*<.01, *** *p*<0.001 (two-tailed tests).

6.3 Motivation to Delay Childbearing

***SSI* did not affect young womens' motivation to delay childbearing.** At both time points, young women in both the treatment and control groups were highly motivated to delay childbearing. Young women in both groups indicated a belief in the importance of delaying childbearing until personal goals have been achieved, and there were no statistically significant differences between the groups.

Exhibit 6.3: Short-Term and Longer-Term Effects of SSI on Motivation to Delay Childbearing

Outcome	Adjusted Treatment Mean[a]	Unadjusted Control Mean	Treatment Effect[b]	SES[c]	*p*-Value
Short-Term Follow-Up					
Motivation to delay childbearing[d]	3.76	3.73	0.03	0.05	.309
Longer-Term Follow-Up					
Motivation to delay childbearing[d]	3.73	3.70	0.03	0.05	.319

Source: Follow-up surveys administered 9 months and 18 months after baseline.

Note: Results in this table are based on 1,805 respondents (short-term survey) and 1,811 respondents (longer-term survey) who provided valid survey responses to relevant items.

[a] The treatment group mean is regression adjusted, calculated as the sum of the control group mean and the regression-adjusted impact estimate (treatment effect).

[b] The treatment effect was estimated in a multi-level model that controls for randomization blocks and other covariates. The treatment effect is expressed in percentage points. Due to rounding, reported treatment effects may differ from differences between reported means for the treatment and control groups.

[c] The SES is the standardized effect size of the difference. For outcomes that are not reported as percentages, the SES is the treatment effect divided by the pooled standard deviation of the treatment and control groups. n/a is not applicable.

[d] Scale score averages responses ranging from 1 to 4. Higher scores indicate higher motivation.

6.4 Intentions

At the short-term follow-up, *SSI* had a large significant program effect on one of the component items, intentions to use a condom during sexual intercourse. A greater percentage of program participants reported that they intended to use a condom during sexual intercourse in the 12 months following the survey compared with participants in the control group (86 percent in the treatment group and 80 percent of the control group). After 18 months, the difference between the groups was no longer statistically significant (80 percent in the treatment group and 77 percent in the control group). *SSI* had no impact on intentions to engage in sexual activity or use birth control during sexual intercourse at either time point.

Exhibit 6.4: Short-Term and Longer-Term Effects of SSI on Intentions

Outcome	Adjusted Treatment Mean[a]	Unadjusted Control Mean	Treatment Effect[b]	*p*-Value
Short-Term Follow-Up				
Sexual intercourse [c]	82.56	83.14	−0.58	.734
Oral sex [c]	65.95	67.05	−1.10	.591
Use a condom if they were to have sexual intercourse [c]	86.31	79.74	6.57***	.000
Use birth control if they were to have sexual intercourse [c]	92.41	91.18	1.23	.357
Longer-Term Follow-Up				
Sexual intercourse [c]	81.36	80.62	0.73	.686
Oral sex[c]	69.66	70.84	−1.18	.562
Use a condom if they were to have sexual intercourse [c]	89.45	88.51	0.94	.542
Use birth control if they were to have sexual intercourse[c]	80.30	77.21	3.09	.124

Source: Follow-up surveys administered 9 months and 18 months after baseline.

Note: Results in this table are based on 1,801–1,804 respondents (short-term survey) and 1,805–1,811 respondents (longer-term survey) who provided valid survey responses to relevant items.

[a] The treatment group mean is regression adjusted, calculated as the sum of the control group mean and the regression-adjusted impact estimate (treatment effect).

[b] The treatment effect was estimated in a multi-level model that controls for randomization blocks and other covariates. The treatment effect is expressed in percentage points. Due to rounding, reported treatment effects may differ from differences between reported means for the treatment and control groups.

[c] Outcomes measure intention to engage in the behavior in the next 12 months. Dichotomous variables, reported as percentage of respondents who responded affirmatively.

6.5 Skills

At both the short-term and longer-term follow-up, *SSI* had a significant positive effect on perceived refusal skills. That is, after nine months, young women in the treatment group were more likely than their control group counterparts to report confidence in their ability to say no to sex. This difference was sustained after 18 months.

There were no differences on perceived ability to successfully negotiate condom use with a partner after nine months. However, at the longer-term follow-up, program participants were more likely than their control group counterparts to report that they could successfully negotiate condom use with a partner.

Exhibit 6.5: Short-Term and Longer-Term Effects of *SSI* on Skills

Outcome	Adjusted Treatment Mean[a]	Unadjusted Control Mean	Treatment Effect[b]	SES[c]	*p*-Value
Short-Term Follow-Up					
Perceived refusal skills [d]	3.45	3.34	0.10***	0.17	.001
Perceived condom negotiation skills [d]	3.73	3.69	0.03	0.08	.126
Longer-Term Follow-Up					
Perceived refusal skills [d]	3.44	3.36	0.07*	0.12	.019
Perceived condom negotiation skills [d]	3.69	3.64	0.05*	0.10	.041

Source: Follow-up surveys administered 9 months and 18 months after baseline.

Note: Results in this table are based on 1,808 respondents (short-term survey) and 1,814–1,815 respondents (longer-term survey) who provided valid survey responses to relevant items.

[a] The treatment group mean is regression adjusted, calculated as the sum of the control group mean and the regression-adjusted impact estimate (treatment effect).

[b] The treatment effect was estimated in a multi-level model that controls for randomization blocks and other covariates. The treatment effect is expressed in percentage points. Due to rounding, reported treatment effects may differ from differences between reported means for the treatment and control groups.

[c] The SES is the standardized effect size of the difference. For outcomes that are not reported as percentages, the SES is the treatment effect divided by the pooled standard deviation of the treatment and control groups. n/a is not applicable.

[d] Scale score averages responses ranging from 1 to 4. Higher scores indicate greater certainty about skills.

* $p<.05$, ** $p<.01$, *** $p<.001$ (two-tailed tests).

7. Exploratory Analyses of Program Effects by Site and Subgroup

The results of the confirmatory analyses reported in Chapter 5 offer the best evidence to answer with confidence whether or not *SSI* had an impact. At the same time, the amount of data collected and pooled across the three sites allowed us to explore other secondary research questions related to possible variation in effects by site or certain subgroups. The results of those analyses, presented in this chapter, must be interpreted with caution and primarily be viewed as hypothesis generating, rather than as additional conclusive evidence on program impacts. The reason for this caution is because the large number of tests conducted in these exploratory analyses increases the risk of producing a significant finding simply by chance, and no adjustments are made to reduce that risk.[22] We cautiously interpret findings in cases where we can identify a pattern of either positive or negative findings in the same direction.

Key Site and Subgroup Findings

Site-Level Differences:

- There were no differences across sites on behavioral outcomes after 18 months. In the short-term, there were differences across sites, with program participants in one site less likely to have engaged in recent oral sex.
- There were some site-level differences in effects on non-behavioral intermediate outcomes at the short-term, but not the longer-term follow-up.

Subgroup Differences

- There were differences on behavioral outcomes at the short-term among subgroups, but there were no effects on behavioral or non-behavioral outcomes after 18 months.

With this caveat, here we present the results of analyses that examined site-level differences in effects on the behavioral outcomes and non-behavioral intermediate outcomes described in the previous chapters. Later sections examine differences in effects on outcomes for different subgroups based on age, race/ethnicity, and baseline sexual experience.

Tables documenting the site-level analyses can be found in Appendix B, and the corresponding tables documenting subgroup analyses are in Appendix C.

7.1 Site-Level Differences

In this section we discuss findings related to site-level differences in effects on both behavioral and non-behavioral intermediate outcomes. We test for site-level differences in effects at both the short-term and longer-term follow-up periods.

7.1.1 Behavioral Outcomes

Exploratory analyses found **no significant site-level differences in program effects after 18 months on behavioral outcomes**, including sexual activity, sexual risk behavior, pregnancy, or contraction of STIs.[23] **In the short-term, there was a significant difference in the effects of *SSI* on engaging in oral sex across sites.** In the Hennepin County site, young women in the treatment group were significantly less likely than their control group counterparts to engage in oral sex in the last 90 days. No other

[22] We would expect to see statistically significant test results for five percent of the tests purely by chance.

[23] For the site-level impact analyses, we conducted a total of 62 tests of the impact of *SSI* on sexual behavior, sexual risk behavior, and the consequences of sexual risk behavior. Ten measures of sexual behavior and sexual risk behavior from the short-term survey, and 12 measures of sexual behaviors, sexual risk behaviors, and consequences from the longer-term survey, at each of two sites (44 measures), plus and nine measures at short-term, and 11 measures at the longer-term in Knox County (no anal sex outcomes), for a total of 62 tests.

treatment-control differences were observed in the Knox County or PPGO sites (Appendices C.1 and C.2).

7.1.2 Non-Behavioral Intermediate Outcomes

At the short-term follow-up, there was a significant difference across sites for intentions to use a condom during sexual intercourse in the subsequent 12 months. In Hennepin County, a significantly greater percentage of program participants reported intentions to use condoms during sexual intercourse in the subsequent 12 months than did control group members (a 10.5 percentage point difference). In Knox County and PPGO, the treatment-control group differences on this outcome were smaller and not statistically significant. At the longer-term follow-up, there were no significant differences in program impact across the sites.

7.2 Subgroup Differences

We also conducted exploratory analyses to look at differences in program effects by subgroups of participants. We specifically looked at whether program effects differed by age, race/ethnicity, and baseline sexual experience. Below we present impact estimates for individual subgroups when there is a statistically significant difference in program impact between subgroups.

7.2.1 Behavioral Outcomes

There were **very few significant differences on the behavioral outcomes by program subgroup**. At the short-term follow-up, we found significant program impacts of *SSI* on the number of lifetime sexual partners for young women who were sexually inexperienced at baseline and for young women who were Hispanic. Compared with their control group counterparts, program youth who were sexually inexperienced were less likely to report having more than one lifetime partner for sexual intercourse.[24] Among Hispanic participants, compared with their control group counterparts, program youth were less likely to report having more than one lifetime partner for sexual intercourse. Also at the short-term follow-up, *SSI* significantly decreased rates of engaging in oral sex without a condom for treatment group members who were older (age 18 and older at baseline) relative to their control group counterparts. There were no program effects on this outcome observed for younger youth (younger than 18 at baseline). The effects on engaging in sexual intercourse in the last 90 days also differed significantly between the two age groups; however, in neither age group did the difference between treatment and control group members reach a conventional level of statistical significance.

7.2.2 Non-Behavioral Intermediate Outcomes

There were **very few subgroup differences on the non-behavioral intermediate outcomes at the short-term, and there were no subgroup differences at the longer-term**. After nine months, there was a significant difference in the impacts of *SSI* on intentions to use a condom during sexual intercourse in the subsequent 12 months for a subgroup defined by sexual experience at baseline. Program participants who had been sexually active at baseline were significantly more likely to express intentions to use condoms during sexual intercourse than were their control group counterparts (Appendix Table C.3). There were no effects on intentions to use condoms during sexual intercourse for those who were sexually

[24] For the subgroup analyses, we conducted 44 tests of variation among subgroups defined by age and sexual experience at baseline (for each of the 10 sexual behavior and sexual risk behavior measures from the short-term survey and 12 measures of sexual behaviors and consequences from the longer-term survey).

inexperienced at baseline. After 18 months, there were no significant differences in intentions based on sexual experience.

8. Discussion

SSI is one of a small number of TPP program models that is designed specifically for females. The aim of the program is to change the sexual risk behavior of young women who are already sexually active.

SSI successfully reduced sexual risk-taking behavior among young women.

The behavioral outcomes identified as key to assessing the effectiveness of TPP programs funded by OAH are more easily measured for this program model, within the timeframe for the study, since vitually all of the young women were sexually active at the time of the intervention and behaving in ways that potentially put them at risk for unplanned pregnancy or STIs or both. Compared with other programs that deliver services to youth many of whom may not be not sexually active at the time of the intervention, we could expect to see behavioral impacts in a relatively short period of time.

Nine months after baseline, young women assigned to the program were less likely to report having unprotected sex (i.e., engaging in sexual intercourse without using some form of birth control)—a 5.8 percentage point difference, or 21 percent fewer. Though the effect diminished over time, this early impact was reflected in the lower pregnancy rates reported by program participants at the longer-term follow-up. The three percentage point difference in pregnancies, though not statistically significant, is practically meaningful because of the long-term consequences and costs associated with unplanned births to teen mothers.

Motivational interviewing is an effective technique for achieving positive change in sexual risk behavior.

The findings also suggest some implications for clinical practice. Motivational interviewing, though successfully used in other clinical practices, has not been previously tested on a large scale in the field of sexual and reproductive health. It seems to have been effective in actively engaging participants and in retaining them (nearly two-thirds of participants attended three out of the four sessions). Faced with the challenge of changing established behaviors, the technique produced changes in the skills and intentions that might ultimately lead to the necessary actions. Although disentangling the roles of motivational interviewing and repeated contact is beyond the scope of this study, it is likely that the tailored follow-up sessions with a health educator supported the change process.

The strong implementation of SSI in a variety of clinical settings and at different levels of scale is unusual and noteworthy.

The selection of *SSI* as a program model for inclusion in the study was driven by interest in reaching young people in clinic settings with effective programs to reduce sexual risk-taking behavior. The broader targeting strategy proposed by all three grantees (to serve young women who were sexually active or contemplating engaging in sexual activity and who were not necessarily seeking medical treatment for an STI) compared with that of the original study raised questions about the feasibility of replicating or expanding *SSI* to achieve similar impacts with a broader population. In reality, in all three replications, *SSI* was implemented with fidelity to the key elements of the intervention. With limited guidance from the program developer (who had tested the program on young women hospitalized for treatment of an STI), each of the grantees independently formulated a strategy to identify, recruit, and retain young women from the community who were sexually active or contemplating be sexually active in the near future.

Most notably, Hennepin County, the largest replication site, implemented the program across a large and diverse set of clinics and managed to impose and maintain uniform standards for program delivery. This counters a common assumption that large-scale replication inevitably means watering down the program or uneven implementation. The experience in Hennepin County suggests that, with sufficient capacity-building support, oversight and continuous monitoring, it is possible to implement *SSI* on a large scale and achieve success in reducing sexual risk behaviors in young women.

The TPP Replication Study was designed to address important research and policy questions about the effectiveness of evidence-based programs and what happens when they are taken to scale and replicated with different populations and in different settings. The three program models (*Safer Sex Intervention (SSI)*, *Reducing the Risk*, and *¡Cuídate!)* were selected to maximize what could be learned about different strategies and to begin to address identified gaps in the teen pregnancy prevention research. This evaluation of *SSI* provides evidence that strong replications of a program for young women can have positive impacts on sexual risk behavior and on subsequent consequences.

References

Bandura, A. (1986). *Social functions of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological), 57*(1), 289-300.

Bloom, H. S. (2006). The core analytics of randomized experiments for social research. MDRC Working Papers on Research Methodology. Retrieved from http://www.mdrc.org/publications/437/full.pdf

Constantine, J., Player, D., Silva, T., Hallgren, K., Grider, M., & Deke, J. (2009). *An evaluation of teachers trained through different routes to certification. Final report.* NCEE 2009-4043. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Firpo-Triplett, R., Rex, P., & Shrier, L. (2011). *Safer Sex Intervention* [Adaptation Kit]. Scotts Valley, CA: ETR Associates.

Gleason, P., Clark, M., Tuttle, C., & Dwyer, E. (2010). *The evaluation of charter school impacts: Executive summary.* NCEE 2010-4030. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Goesling, B., Colman, C., Trenholm, C., Terzian, M., & Moore, K. (2014). Programs to reduce teen pregnancy, sexually transmitted infections, and associated risk behaviors: A systematic review. *Journal of Adolescent Health, 54*(5), 499-507.

Kappeler, E., & Farb, A. (2014). Historical context for the creation of the Office of Adolescent Health and the Teen Pregnancy Prevention Program. *Journal of Adolescent Health, 54*(3S), S3-S9.

Kelsey, M., Walker, J. T., Layzer, J., Price, C., & Juras, R. (2016). Replicating the Safer Sex Intervention: 9-month impact findings of a randomized controlled trial. *American Journal of Public Health, 106*(S1), S53-S59.

Orr, L. L. (1999). *Social experiments: Evaluating public programs with experimental methods.* Thousand Oaks, CA: Sage Publications.

Prochaska, J. O., & DiClemente, C. C. (2005). The transtheoretical approach. In J. C. Norcross & M. R. Goldfried (Eds.), *Handbook of psychotherapy integration* (2nd ed., pp. 147-171). New York: Oxford University Press.

Prochaska, J. O., & Velicer, W. F. (1997). The transtheoretical model of health behavior change. *American Journal of Health Promotion, 12*(1), 38-48.

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.

Schochet P.Z. (2008a). *Technical Methods Report: Guidelines for Multiple Testing in Impact Evaluations* (NCEE 2008-4018). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Schochet, P. Z. (2008b). Statistical power for random assignment evaluations of education programs. *Journal of Educational and Behavioral Statistics, 33*(1), 62-87.

Shrier, L. A., Ancheta, R., Goodman E., Chiou, V. M., Lyden, M. R., & Emans, S. (2001). Randomized controlled trial of a safer sex intervention for high-risk adolescent girls. *Archives of Pediatrics & Adolescent Medicine, 155*(1), 73-79.

U.S. Department of Health and Human Services [HHS]. (2010). *Teenage Pregnancy Prevention: Replication of evidence-based programs. Funding opportunity announcement and application instructions.* Washington, DC: Office of Adolescent Health and Office of Public Health and Science, U.S. Department of Health and Human Services.

Appendix A: Measures

The measures we used to examine short-term and longer-term program impacts stem from our research questions (Section 3.1) and logic model (Exhibit 2.3) and are organized into two categories:

- Youth sexual activity, sexual risk behavior, and consequences of sexual risk behavior; and
- Non-behavioral intermediate outcomes.

Measures of *youth sexual activity, sexual risk behavior,* and *consequences of sexual risk behavior* include recent sexual activity, sexual intercourse, oral sex, anal sex, recent sexual intercourse without birth control, sexual intercourse without a condom, oral sex without a condom, anal sex without a condom, and pregnancy and STI diagnoses. Measures of *non-behavioral intermediate outcomes* indicate the extent to which youth assimilated the program's messages and reflected them in their knowledge, attitudes, motivation, intentions, and skills—all of which are hypothesized precursors of change in youth's sexual behavior. In the sections that follow, we describe each category by defining constituent measures and their construction.

A.1 Youth Sexual Activity, Sexual Risk Behavior, and Consequences of Sexual Risk Behavior

To understand program effects on youths' sexual activity, sexual risk behavior and consequences of sexual risk behavior, we examined the 13 items presented in Exhibit A.1.[25]

Exhibit A.1: Youth Sexual Behavior, Sexual Risk Behavior, and Sexual Consequences Measures

Measure	Item	Coding
Sexual Behavior Outcomes		
Sexual Activity		
Recent sexual activity (in the last 90 days)	Coded from three separate items measuring sexual intercourse in the last 90 days, oral sex in the last 90 days, and anal sex in the last 90 days.	Youth who reported they had engaged in one or more of the sexual activities (sexual intercourse, oral sex, or anal sex) during the last 90 days received a score of 1 on this measure. Youth who reported no sexual activity during the last 90 days received a score of 0, as did those who reported (on a separate question) that they had never been sexually active. Note that *sexual activity* is defined differently across grantees. In Hennepin County and Planned Parenthood of Greater Orlando, sexual activity refers to sexual intercourse, oral sex, and anal sex. Youth were not asked about anal sex in Knox County.
Sexual intercourse in the last 90 days	Now please think about the past 3 months. In the past 3 months, have you had sexual intercourse?	Youth responded to this question with a yes(1)/no(0) answer. Youth who reported engaging in sexual intercourse in the last 90 days received a score of 1 on the measure. Those who reported they had not engaged in sexual intercourse in the last 90 days received a score of 0 on the measure, as did those who reported (on a separate question) that they had never been sexually active.

[25] Note that 12 outcomes are reported in the main results presented in Section 4. The 13th outcome, "initiation of sexual activity," appears only in Appendix A because it is relevant only to the small subgroup of youth who were sexually inexperienced at baseline. In Appendix E, we present the impacts on sexual behavior and sexual risk among this subgroup alone.

Measure	Item	Coding
Oral sex in the last 90 days	Now please think about the past 3 months. In the past 3 months, have you had oral sex?	Youth responded to this question with a yes(1)/no(0) answer. Youth who reported engaging in oral sex in the last 90 days received a score of 1 on the measure. Those who reported they had not engaged in oral sex in the last 90 days received a score of 0 on the measure, as did those who reported (on a separate question) that they had never been sexually active.
Anal sex in the last 90 days	Now please think about the past 3 months. In the past 3 months, have you had anal sex?	Youth responded to this question with a yes(1)/no(0) answer. Youth who reported engaging in anal sex in the last 90 days received a score of 1 on the measure. Those who reported they had not engaged in anal sex in the last 90 days received a score of 0 on the measure, as did those who reported (on a separate question) that they had never been sexually active. Note that youth were not asked about anal sex in Knox County.
Initiation of sexual activity	Have you ever had any of the following: sexual intercourse, oral sex, or anal sex?	Youth responded to this question with a yes(1)/no(0) answer. This item was coded 0 or 1, with 1 representing one or more forms of sexual activity (sexual intercourse, oral sex, and/or anal sex) during one's lifetime and 0 representing no sexual activity during one's lifetime. Responses to other sexual behavior and sexual risk questions were examined and back-coded into this question such that youth who reported they had engaged in one or more of the sexual activities received a score of 1. Note that *sexual activity* is defined differently across grantees. In Hennepin County and Planned Parenthood of Greater Orlando, sexual activity refers to sexual intercourse, oral sex, and anal sex. Youth were not asked about anal sex in Knox County.
Sexual Risk Behavior		
Sexual intercourse without birth control (in the last 90 days)	In the past 3 months, have you had sexual intercourse without you or your partner using any of these methods of birth control, even just once? • Condoms • Birth control pills • The shot (Depo-Provera) • The patch • The ring (NuvaRing) • IUD (Mirena or Paragard) • Implants (Implanon)	Youth responded to this question with a yes(1)/no(0) answer. Youth who reported engaging in sexual intercourse without birth control in the last 90 days received a score of 1 on the measure. Those who reported they had *not* engaged in sexual intercourse without birth control in the last 90 days received a score of 0 on the measure, as did those who reported (on a separate question) that they had never been sexually active.
Sexual intercourse without a condom (in the last 90 days)	In the past 3 months, have you had sexual intercourse without your partner using a condom?	Youth responded to this question with a yes(1)/no(0) answer. Youth who reported engaging in sexual intercourse without a condom in the last 90 days received a score of 1 on the measure. Those who reported they had *not* engaged in sexual intercourse without a condom in the last 90 days received a score of 0 on the measure, as did those who reported (on a separate question) that they had never been sexually active.
Oral sex without a condom (in the last 90 days)	In the past 3 months, have you had oral sex without using a condom, even once?	Youth responded to this question with a yes(1)/no(0) answer. Youth who reported engaging in oral sex without a condom in the last 90 days received a score of 1 on the measure. Those who reported they had *not* engaged in oral sex without a condom in the last 90 days received a score of 0 on the measure, as did those who reported (on a separate question) that they had never been sexually active.

Measure	Item	Coding
Anal sex without a condom (in the last 90 days)	In the past 3 months, have you had anal sex without using a condom, even once?	Youth responded to this question with a yes(1)/no(0) answer. Youth who reported engaging in anal sex without a condom in the last 90 days received a score of 1 on the measure. Those who reported they had *not* engaged in anal sex without a condom in the last 90 days received a score of 0 on the measure, as did those who reported (on a separate question) that they had never been sexually active. Note that youth were not asked about anal sex in Knox County.
Sexual intercourse with more than one partner (lifetime)	How many different people have you ever had sexual intercourse with, even if only one time?	Youth responded to this question on a scale from 0 to 100. This item was coded 0 or 1, with 1 representing multiple sexual partners and 0 representing one or no sexual partners in one's lifetime.
Sexual intercourse with more than five partners (lifetime)	How many different people have you ever had sexual intercourse with, even if only one time?	Youth responded to this question on a scale from 0 to 100. This item was coded 0 or 1, with 1 representing six or more sexual partners and 0 representing five or fewer (including zero) sexual partners in one's lifetime.
Sexual Consequences (Longer-term follow-up only)		
Pregnant since baseline	To the best of your knowledge, have you ever been pregnant, even if no baby was born?	This outcome measure was coded as 1=yes, 0=no indicating whether or not respondents reported that they had been pregnant between baseline and the longer-term follow up. When youth reported a greater number of pregnancies at the longer-term survey than at baseline, the youth was assigned a score of 1. Youth who reported the same number at baseline and the longer-term follow-up were assigned a score of 0.
Diagnosed with STI in the last 12 months	In the past 12 months, have you been told by a doctor or nurse that you had a sexually transmitted disease (STD) / sexually transmitted infection (STI) or HIV?	Youth responded to this question with a yes(1)/no(0) answer.

A.2 Non-Behavioral Intermediate Outcomes

Non-behavioral intermediate outcomes are those expected to portend changes in behavior. We asked youth a wide variety of questions to gauge their understanding, thoughts, beliefs, and perceptions of topics addressed by the program. We organized these measures conceptually into five domains: knowledge, attitudes, motivation, intentions, and skills. Using survey items relevant to each domain, we conducted factor analyses and reliability testing to construct composite measures in each domain, where this was possible. In addition, we used baseline data (when the same items were asked) to examine the stability over time of composite measures, and examined the follow-up data by racial/ethnic subgroup to assess the stability of constructs.

Knowledge

To examine program-related changes in youth's sexual health knowledge, we constructed two measures: *knowledge of pregnancy risk* and *knowledge of STI risk.* These measures were defined conceptually and constructed to differentiate accurate knowledge from misinformation. They may be considered tests of understanding of the factors contributing to pregnancy and STIs. The construction of these measures is described below and detailed information about their component items is presented in A.2

- ***Knowledge of pregnancy risk*** is a composite measure that is the mean (multiplied by 100) of four binary variables regarding knowledge of the extent to which contraceptive methods can prevent

pregnancy and circumstances under which pregnancy is possible. (See Exhibit A.2 for coding and other details.) Scores on this scale range from 0 to 100 and represent the percentage of correct answers across the four items. Higher values indicate more accurate knowledge.

- ***Knowledge of STI risk*** is a composite measure that is the mean (multiplied by 100) of 12 binary variables pertaining to knowledge of STI prevention, transmission, and treatment. (See Exhibit A.2 for coding and other details.) Scores on this scale range from 0 to 100 and represent the percentage of correct answers across the 12 items. Higher values indicate more accurate knowledge.

Exhibit A.2: Knowledge Scales and Component Items

Component Items	Coding
Knowledge of Pregnancy Risk (4 items)	
Used correctly, how much can birth control pills reduce pregnancy risk?	Youth responded to this question on a scale from 1="Not at all" to 4="Completely." This item was recoded into a binary variable where the correct response ("A lot") was coded as 1 and all other responses were coded as 0.
Used correctly, how much can condoms reduce pregnancy risk?	
A couple that has had unprotected sex and not gotten pregnant does not have to worry about getting pregnant.	Youth indicated the veracity of this statement, responding on a scale from 1="I am sure it's true" to 5="I am sure it's false." This item was recoded into a binary variable where 1 indicates youth were sure or thought the statement was false, and 0 indicates they were sure or thought the statement was true or did not know.
A woman is protected from pregnancy the day she begins taking the pill.	
Knowledge of STI Risk (12 items)	
You can't get infected with HIV if you have sex only once or twice without a condom.	Youth indicated the veracity of this statement, responding on a scale from 1= "I am sure it's true" to 5 = "I am sure it's false." This item was recoded into a binary variable where 1 indicates youth were sure or thought the statement was false and 0 indicates they were sure or thought the statement was true or did not know.
Once you are infected with HIV you are infected for life.	
There is a vaccine to prevent girls from getting HPV.	
All STDs/STIs can be cured by taking medicine.	
A person with an STD/STI who looks and feels healthy cannot transmit the infection to others.	
Some STDs/STIs put you at greater risk of HIV.	
About 1 out of 4 sexually active teens gets an STD/STI every year.	
You can get an STD/STI from having oral sex.	
Used correctly, how much can condoms decrease the risk of HIV?	Youth responded to this question on a scale from 1="Not at all" to 4="Completely." This item was recoded into a binary variable where the correct response ("not at all") was coded as 1 and all other responses were coded as 0.
Used correctly, how much can condoms decrease the risk of gonorrhea?	
Used correctly, how much can birth control pills decrease the risk of HIV?	
Used correctly, how much can birth control pills decrease the risk of gonorrhea?	

Attitudes

The short-term survey included 24 items querying attitudes toward sexual behaviors, sexual risks, and contraceptive methods. From among these, we constructed two measures to examine program impacts on youths' sexual health attitudes: *attitudes toward protection* and *attitudes toward risky sexual behavior*. These measures are described below and detailed information about their component items is presented in Exhibit A.3.

- ***Attitudes toward protection*** is a composite measure that is the mean of responses to 12 items about the importance of using condoms and/or birth control during sexual activity. (See Exhibit A.3 for coding and other details.) Scores on this scale represent the level of support for using protection. They range from 1 to 4, with high scores indicating positive and supportive attitudes toward contraceptive use to prevent STIs and/or pregnancy. The measure demonstrated acceptable internal consistency reliability ($\alpha = 0.78$).[26]
- ***Attitudes toward risky sexual behavior*** is a composite measure that is the mean of seven binary items (multiplied by 100) querying the acceptability and normativeness of extremely risky sexual behaviors. (See Exhibit A.3 for coding and other details.) Scores on this scale range from 0 to 100 and represent the percentage of items agreed with. Higher values reflect more support for risky behavior. The measure demonstrated acceptable internal consistency reliability ($\alpha = 0.79$).

Exhibit A.3: Attitudes Scales and Component Items

Component Items	Coding
Attitudes Toward Protection (12 items)	
Birth control pills should always be used if a person your age has sexual intercourse.	Youth expressed their agreement with this statement, responding on a scale from 1="Strongly agree" to 4="Strongly disagree." High values indicate more positive attitudes toward condoms.
Birth control is too much trouble to use.	
Birth control is pretty easy to get.	
Birth control is important to make sex safer.	
Birth control has too many side effects.	
Using birth control is morally wrong.	
Condoms are too much trouble to use.	
Condoms are pretty easy to get.	
Condoms are important to make sex safer.	
Using condoms means you don't trust your partner.	
Using condoms is morally wrong.	
Condoms decrease sexual pleasure.	
Attitudes Toward Risky Sexual Behavior (7 items)	
It's OK to have sex with someone on your first date.	Youth expressed their agreement with this statement by selecting it if it reflected their views on engaging in sex. Responses were coded in a binary fashion, as 1 when the statement was selected and 0 when not selected.
It's OK to have sex with someone the same night you meet them.	
It's OK to have sex with several different people in the same month.	
It's OK to have sex without protection.	
It's OK to have sex with someone when you know they are someone else's girlfriend/boyfriend.	
It's OK to have sex with someone if you are drunk or high.	
It's OK to have sex with someone if you know they are drunk or high.	

[26] As a general rule of thumb, the internal validity of scales with reliability coefficients of 0.70–0.79 is considered "acceptable," of 0.80–0.89 is considered "good," and 0.90 or greater is considered "excellent."

Motivation

The short-term and longer-term surveys included 22 items related to youth's motivation to engage in safe sexual practices and reduce their risk. From these, we developed a measure of *motivation to delay childbearing*. It is the average of three items related to reasons for delaying childbearing. (See Appendix Exhibit A.4 for coding and other details.) Scores on this scale range from 1 to 4 with higher scores indicating more motivation to wait to have a child. The scale demonstrated good internal consistency reliability ($\alpha = 0.88$).

Exhibit A.4: Motivation Scale and Component Items

Component Items	Coding
Motivation to Delay Childbearing (3 items)	
You have goals you want to accomplish before having a child.	
It is important for you to finish school before you have a child.	Youth responded to this question on a scale from 1="Strongly agree" to 4="Strongly disagree." We reverse-coded this item so that higher values indicate more agreement.
It is important to have a job and a stable income before you have a child.	

Intentions

We used the four items presented in Exhibit A.5 to examine impacts on youth's intended or anticipated sexual behavior and sexual risk behavior in the coming year.

Exhibit A.5: Intentions Measures

Item	Coding
Do you intend to have sexual intercourse in the next year, if you have the chance?	Youth responded to this question on a scale from 1="Yes, definitely" to 4="No, definitely not." This item was recoded into a binary variable where affirmative responses (definitely, probably) were coded as 1 and negative responses (definitely not, probably not) were coded as 0.
Do you intend to have oral sex in the next year, if you have the chance?	
If you have sexual intercourse in the next year, do you intend to use birth control?	
If you have sexual intercourse in the next year, do you intend to use a condom?	

Skills

The short-term and longer-term follow-up surveys included items regarding skills important to reproductive health. From these, we constructed two measures to examine program impacts on youth's perceived ability say no to sex (*refusal skills*) and successfully negotiate condom use with a partner (*condom negotiation skills*). These measures are described below and detailed information about their component items is presented in Exhibit A.6.

- ***Refusal skills*** is a composite measure that is the mean of responses to six items about perceived ability to say no to sex in a variety of situations. (See Exhibit A.6 for coding and other details.) Scores on this scale range from 1 to 4, with high scores indicating more confidence in one's abilities to abstain from intercourse. The measure demonstrated good internal consistency reliability ($\alpha = 0.83$).
- ***Condom negotiation skills*** is a composite measure that is the mean of responses to seven items about perceived ability to obtain and negotiate the use of condoms. (See Exhibit A.6 for coding and other details.) Scores on this scale range from 1 to 4, with high scores indicating more confidence in one's abilities to obtain and negotiate the use of condoms. The measure demonstrated good internal consistency reliability ($\alpha = 0.84$).

Exhibit A.6: Skills Scales and Component Items

Component Items	Coding
Refusal Skills (6 items)	
How sure are you that you would be able to say no to having sexual intercourse if your partner really wanted to, but you were not ready?	
How sure are you that you would be able to say no to having sexual intercourse if you just met someone you really liked and that person wanted to have sex, but you didn't?	Youth responded to this question on a scale from 1="I'm sure I could" to 4="I'm sure I could not." We reverse coded this item so that higher values indicate more confidence in one's ability.
How sure are you that you would be able to say no to having sexual intercourse if you had strong sexual feelings for that person?	
How sure are you that you would be able to say no to having sexual intercourse if neither you nor your partner had any form of birth control?	
How sure are you that you would be able to say no to having sexual intercourse if you have dated for a long time?	
How sure are you that you would be able to say no to having sexual intercourse after you have been drinking alcohol?	
Condom Negotiation Skills (7 items)	
If you were going to have sex could you get or buy a condom?	
If you were going to have sex could you talk about using condoms with your partner before having sex?	
If you were going to have sex could you insist on using a condom if your partner didn't want to use one?	
If you were going to have sex could you ask your partner to use condoms even if the two of you had sex before without using condoms?	Youth responded to this question on a scale from 1="I'm sure I could" to 4="I'm sure I could not." We reverse coded this item so that higher values indicate more confidence in one's ability.
If you were going to have sex could you use a condom without spoiling the mood?	
If you were going to have sex could you ask a new partner to use condoms?	
If you were going to have sex could you get a partner to use condoms, even if you're drunk or high?	

Appendix B: Site-Level Effects

This study was carefully designed such that when data from all three replication sites were pooled into a single analysis, the combined sample would be large enough for the study to be adequately powered to detect effects of the *Safer Sex Intervention* on all of the outcomes of interest. Although the pooled analysis is the primary focus of this study, there was clearly considerable interest on the part of study stakeholders in examining the results from each of the three replication sites, and the large sample sizes preserve the ability to conduct these analyses. Therefore, this appendix presents site-specific impact estimates for each of the outcomes reported in the main text.

We urge two major types of caution for readers who examine the results from the individual sites. The first is that the study was not designed to have large enough sample sizes in each individual site to have a good chance of detecting a treatment effect for all of the outcomes of interest. Thus, in a single site, lack of statistical significance could be the result of an insufficiently large sample to detect a true effect, or it could mean that the intervention did not produce an effect on the outcome. Second, there are a large number of results presented in Appendix B, and these results are not adjusted for multiple comparisons. Some statistically significant findings would be expected purely by chance among such a large number of tests. Therefore, the findings in these tables should be interpreted with caution. The final column of each table shows the statistical result for a test of differences in the treatment effect across sites. When a statistically significant difference is found, the corresponding site-specific effects are discussed in the main text, as we only interpret site-specific effects when a significant difference across sites is found.

Exhibit B.1 Short-Term Effects on Sexual Activity and Sexual Risk Behavior by Site

Outcome	Hennepin County (n = 963)				Knox County (n = 412)				Planned Parenthood of Greater Orlando (n = 426)				*p*-Value for the Test of Differences Across Sites[a]
	Adj. Treatment Mean[b]	Unadj. Control Mean	Treatment Effect[c]	*p*-Value	Adj. Treatment Mean[b]	Unadj. Control Mean	Treatment Effect[c]	*p*-Value	Adj. Treatment Mean[b]	Unadj. Control Mean	Treatment Effect[c]	*p*-Value	
Sexual activity (percentage responding affirmatively) [d]													
Recently sexually active (in last 90 days)	76.19	78.66	−2.47	.361	74.03	68.61	5.42	.188	72.41	72.60	−0.19	.963	.277
Sexual intercourse in the last 90 days	72.92	76.22	−3.30	.236	70.74	66.42	4.32	.308	67.93	68.49	−0.56	.893	.322
Oral sex in the last 90 days	57.97	64.02	−6.05 *	.044	61.22	55.47	5.75	.209	60.30	56.85	3.45	.439	.050*
Anal sex in the last 90 days	9.59	5.50	4.09 *	.027	n/a	n/a	n/a	n/a	8.10	7.53	0.57	.835	.287
Sexual risk behavior (percentage responding affirmatively)													
Sexual intercourse without birth control (in last 90 days)	21.43	29.27	−7.84 **	.005	17.58	23.36	−5.78	.173	27.57	28.77	−1.20	.772	.412
Sexual intercourse without a condom (in last 90 days)	57.43	62.80	−5.37	.077	49.94	53.28	−3.34	.470	48.60	49.32	−0.72	.873	.689
Oral sex without a condom (in last 90 days)	53.20	60.06	−6.86 *	.024	57.63	54.74	2.89	.533	53.49	50.68	2.81	.535	.092
Anal sex without a condom (in last 90 days)	7.75	4.59	3.16	.060	n/a	n/a	n/a	n/a	6.38	4.79	1.59	.524	.601
Sexual intercourse with more than one partner (lifetime)	68.69	71.16	−2.47	.316	69.98	70.31	−0.33	.931	73.17	74.65	−1.48	.688	.891
Sexual intercourse with more than five partners (lifetime)	25.85	30.09	−4.24	.085	29.55	29.69	−0.14	.971	24.49	25.35	−0.86	.814	.581

Source: Follow-up survey administered nine months after baseline.
Note: na is not asked.
[a] This column shows the results for statistical tests of whether the treatment effect varies among the three sites.
[b] The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression-adjusted impact estimate (treatment effect).
[c] The treatment effect was estimated in a regression model that controls for randomization blocks and other covariates. The treatment effect is expressed as a difference in percentage points. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.
[d] *Sexual activity* is defined differently across grantees. In Hennepin County and Planned Parenthood of Greater Orlando, sexual activity refers to sexual intercourse, oral sex, and/or anal sex. Youth were not asked about anal sex in Knox County.
* $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed tests).

Exhibit B.2 Longer-Term Impacts on Sexual Activity, Sexual Risk Behavior, and Sexual Consequences by Site

	Hennepin County (n = 952)				Knox County (n = 428)				Planned Parenthood of Greater Orlando (n = 428)				*p*-Value for the Test of Differences Across Sites[a]
	Adj. Treatment Mean[b]	Unadj. Control Mean	Treatment Effect[c]	*p*-Value	Adj. Treatment Mean[b]	Unadj. Control Mean	Treatment Effect[c]	*p*-Value	Adj. Treatment Mean[b]	Unadj. Control Mean	Treatment Effect[c]	*p*-Value	
Sexual Behavior													
Sexual activity (percentage responding affirmatively) [d]													
Currently sexually active (in last 90 days)	75.28	79.11	−3.83	.171	78.76	74.15	4.61	.259	71.03	71.53	−0.50	.904	.231
Sexual intercourse in the last 90 days	72.35	75.63	−3.28	.252	76.35	73.47	2.88	.492	66.09	64.58	1.51	.720	.402
Oral sex in the last 90 days	61.49	63.29	−1.80	.560	64.74	59.86	4.88	.280	54.43	58.33	−3.90	.390	.341
Anal sex in the last 90 days	9.29	10.44	−1.15	.560	n/a	n/a	n/a	n/a	8.78	9.03	−0.25	.932	.796
Sexual risk behavior (percentage responding affirmatively)													
Sexual intercourse without birth control (in last 90 days)	22.65	23.73	−1.08	.712	27.24	32.65	−5.41	.208	22.98	27.08	−4.10	.342	.669
Sexual intercourse without a condom (in last 90 days)	58.54	63.61	−5.07	.114	55.94	61.90	−5.96	.203	48.12	45.83	2.29	.626	.361
Oral sex without a condom (in last 90 days)	57.55	59.49	−1.94	.535	61.78	58.50	3.28	.473	47.70	52.78	−5.08	.268	.422
Anal sex without a condom (in last 90 days)	6.98	9.81	−2.83	.114	n/a	n/a	n/a	n/a	6.39	5.56	0.83	.751	.248
Sexual intercourse with more than one partner (lifetime)	73.36	75.00	−1.64	.535	76.14	71.53	4.61	.235	74.46	72.92	1.54	.690	.397
Sexual intercourse with more than five partners (lifetime)	33.79	33.65	0.14	0.961	35.82	31.25	4.57	.256	27.03	26.39	0.64	.872	.647
Sexual consequences (percentage responding affirmatively)													
Pregnant since baseline	14.53	19.27	−4.74	.070	19.69	25.35	−5.66	.133	15.30	13.43	1.87	.629	.287
Diagnosed with STI in the last 12 months	10.69	14.51	−3.82	.063	7.50	7.48	0.02	.995	9.44	6.94	2.50	.406	.191

Source: Follow-up survey administered 18 months after baseline.
Note: na is not asked.
[a] This column shows the results for statistical tests of whether the treatment effect varies among the three sites.
[b] The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression-adjusted impact estimate (treatment effect).
[c] The treatment effect was estimated in a regression model that controls for randomization blocks and other covariates. The treatment effect is expressed as a difference in percentage points.
[d] *Sexual activity* is defined differently across grantees. In Hennepin County and Planned Parenthood of Greater Orlando, sexual activity refers to sexual intercourse, oral sex, and anal sex. Youth were not asked about anal sex in Knox County.
* $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed tests).

Exhibit B.3 Short-Term Effects on Non-Behavioral Intermediate Outcomes by Site

Outcome	Hennepin County (n = 968)					Knox County (n = 413)					Planned Parenthood of Greater Orlando (n = 427)					*p*-Value for the Test of Differences Across Sites[a]
	Adj. T Mean[b]	Unadj. C Mean	T Effect[c]	*p*-Value	SES[d]	Adj. T Mean[b]	Unadj. C Mean	T Effect[c]	*p*-Value	SES[d]	Adj. T Mean[b]	Unadj. C Mean	T Effect[c]	*p*-Value	SES[d]	
Knowledge																
Knowledge of pregnancy risk [e]	78.91	77.81	1.10	.485		80.61	83.33	−2.72	.258		75.76	74.49	1.27	.589		.367
Knowledge of STI risk [e]	75.62	75.35	0.27	.811		77.46	75.30	2.16	.217		75.04	73.06	1.98	.245		.561
Attitudes																
Attitudes toward protection [f]	3.36	3.34	0.02	.520	0.04	3.35	3.34	0.01	.740	0.03	3.37	3.27	0.10**	.006	0.25	.125
Attitudes toward risky sexual behavior [g]	3.92	5.64	−1.72	.070		3.18	4.90	−1.72	.236		5.42	5.38	0.04	.978		.555
Motivation																
Motivation to delay childbearing[f]	3.75	3.73	0.03	.476	0.05	3.76	3.73	0.03	.620	0.05	3.76	3.74	0.03	.594	0.05	.999
Intentions (Percentage of respondents reporting intentions to engage in the following behaviors in the next 12 months)[h]																
Sexual intercourse	84.57	85.98	−1.41	.549		78.59	81.75	−3.16	.380		81.79	78.08	3.71	.289		.342
Oral sex	65.64	68.39	−2.75	.326		66.42	66.67	−0.25	.954		66.15	64.38	1.77	.671		.650
Use a condom if they were to have sexual intercourse	86.82	76.29	10.53 ***	.000		81.23	83.21	−1.98	.592		90.14	84.25	5.89	.102		.018*
Use birth control if they were to have sexual intercourse	94.00	91.19	2.81	.126		92.87	90.51	2.36	.398		88.45	91.78	−3.33	.222		.156
Skills																
Perceived refusal skills (scale score) [e]	3.43	3.36	0.08	.064	0.13	3.45	3.36	0.10	.123	0.16	3.46	3.30	0.17 **	.007	0.27	.470
Perceived condom negotiation skills (scale score) [f]	3.73	3.70	0.03	.318	0.07	3.74	3.69	0.05	.239	0.12	3.72	3.70	0.02	.623	0.05	.870

Source: Follow-up survey administered nine months after baseline.

[a] This column shows the results for statistical tests of whether the treatment effect varies among the three sites.

[b] The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression adjusted impact estimate (treatment effect).

[c] The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. For outcomes reported as percentages, the treatment effect is expressed as a difference in percentage points. For scale outcomes, the treatment effect is expressed in the original metric of the outcome variable. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.

[d] The effect size is the standardized effect size of the difference, which is the "treatment effect" divided by the pooled standard deviation of the treatment and control groups.

[e] Scores represent the average percent of items answered correctly.

[f] Scale score averages responses ranging from 1 to 4. Higher scores indicate higher levels of the outcome.

[g] Score represents the average percentage of items agreed with.

[h] Dichotomous variables, reported as percentage of respondents who responded affirmatively.

* $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed tests).

Exhibit B.4 Longer-Term Effects on Non-Behavioral Intermediate Outcomes by Site

Outcome	Hennepin County (n = 958)					Knox County (n = 429)					Planned Parenthood of Greater Orlando (n = 428)					*p*-Value for the Test of Differences Across Sites[a]
	Adj. T Mean[b]	Unadj. C Mean	T Effect[c]	*p*-Value	SES[d]	Adj. T Mean[b]	Unadj. C Mean	T Effect[c]	*p*-Value	SES[d]	Adj. T Mean[b]	Unadj. C Mean	T Effect[c]	*p*-Value	SES[d]	
Knowledge																
Knowledge of pregnancy risk [e]	77.40	78.37	−0.97	.569		81.02	81.25	−0.23	.927		73.97	73.96	0.01	.997		.938
Knowledge of STI risk [e]	76.16	76.44	−0.28	.822		76.15	76.01	0.14	.936		72.93	73.96	−1.03	.572		.897
Attitudes																
Attitudes toward protection [f]	3.32	3.31	0.01	.572	0.04	3.33	3.31	0.02	.578	0.05	3.31	3.24	0.07	.088	0.16	.535
Attitudes toward risky sexual behavior [g]	5.17	6.49	−1.32	.200	0.01	3.20	6.12	−2.92	.054	−0.07	6.40	7.44	−1.04	.492	0.06	.616
Motivation																
Motivation to delay childbearing[f]	3.73	3.74	−0.01	.847	−0.01	3.74	3.66	0.08	.146	0.15	3.71	3.66	0.05	.383	0.09	.385
Intentions to engage in the following behaviors in the next 12 months																
Sexual intercourse[h]	82.45	83.60	−1.15	.646		79.34	77.03	2.31	.528		80.99	77.78	3.21	.384		.547
Oral sex[h]	69.80	72.78	−2.98	.293		69.71	68.03	1.68	.684		69.23	69.44	−0.21	.959		.625
Use a condom if they were to have sexual intercourse[h]	80.09	75.47	4.62	.097		79.51	78.38	1.13	.780		81.62	79.86	1.76	.667		.725
Use birth control if they were to have sexual intercourse[h]	90.02	90.85	−0.83	.697		91.00	89.19	1.81	.561		86.53	82.64	3.89	.215		.437
Skills																
Perceived refusal skills [f]	3.44	3.40	0.05	.284	0.07	3.47	3.36	0.11	.081	0.18	3.39	3.30	0.09	.142	0.15	.657
Perceived condom negotiation skills [f]	3.68	3.66	0.02	.495	0.05	3.72	3.66	0.06	.199	0.13	3.69	3.59	0.09	.058	0.19	.469

Source: Follow-up survey administered 18 months after baseline.

[a] This column shows the results for statistical tests of whether the treatment effect varies among the three sites.

[b] The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression adjusted impact estimate (treatment effect).

[c] The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. For outcomes reported as percentages, the treatment effect is expressed as a difference in percentage points. For scale outcomes, the treatment effect is expressed in the original metric of the outcome variable. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.

[d] The effect size is the standardized effect size of the difference, which is the "treatment effect" divided by the pooled standard deviation of the treatment and control groups.

[e] Scores represent the average percent of items answered correctly.

[f] Scale score averages responses ranging from 1 to 4. Higher scores indicate higher levels of the outcome.

[g] Score represents the average percentage of items agreed with.

[h] Dichotomous variables, reported as percentage of respondents who responded affirmatively.

* $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed tests).

Appendix C: Subgroup Effects

To better understand what works for whom, we estimated effects for key subgroups of participants (based on age, race/ethnicity, and sexual experience at baseline) and tested for differences in effects between subgroups. To guard against potential overinterpretation of results, we present impact estimates for individual subgroups only when there is a statistically significant difference between subgroups. For example, the impact estimate would be presented for the subgroup of Hispanic respondents only if there were a statistically significant difference between the effects on respondents across race/ethnicity.

Exhibit C.1: Short-Term Effects on Sexual Activity and Sexual Risk Behavior, by Subgroup

	Treatment Effect[a]	*p*-Value[b]
Sexual intercourse in the last 90 days		
Subgroup: Respondent age		
Respondent less than age 18 (*n* = 962)	2.96	.294
Respondent age 18 or older (*n* = 839)	−5.52	.060
Oral sex without a condom in the last 90 days		
Subgroup: Respondent age		
Respondent less than age 18 (*n* = 963)	2.05	.507
Respondent age 18 or older (*n* = 838)	−7.43*	.021
Sexual intercourse with more than one lifetime sexual partner		
Subgroup: Respondent sexual experience at baseline		
Never sexually active at baseline (*n* = 115)	−21.42*	.028
Ever sexually active at baseline (*n* = 1,620)	−0.50	.790
Subgroup: Respondent race/ethnicity[c]		
Hispanic (*n* = 309)	−9.75*	.020
Black (*n* = 612)	1.03	.736
White (*n* = 580)	2.53	.422
Other (*n* = 234)	−8.31	.097

Source: Follow-up survey administered nine months after baseline.
Notes: Impact estimates for subgroups are shown only if a test for differences in impacts among the subgroups met the study criterion for statistical significance ($p < .05$). For example, a test result indicated that the treatment effect on more than one lifetime partner (for sexual intercourse) was significantly different between the sexually experienced at baseline subgroups.

[a] This column shows the estimated treatment effect (treatment/control difference in percent responding affirmatively) for the subgroup indicated in the row.

[b] This column shows the statistical test result for whether the treatment effect for the subgroup indicated in the row was significantly different than zero.

[c] Racial-ethnic categories are Hispanic, Black non-Hispanic, White non-Hispanic, and Other race non-Hispanic, where Other is defined as Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, multiracial, or undisclosed race.

* $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed tests).

Exhibit C.2 Longer-Term Effects on Sexual Activity and Sexual Risk Behavior by Subgroup

	Treatment Effect [a]	*p*-Value [b]
More than one lifetime sexual partner (for sexual intercourse)		
Subgroup: Respondent age		
Respondent less than age 18 (*n* = 957)	4.65	.080
Respondent age 18 or older (*n* = 831)	−3.97	.148

Source: Follow-up survey administered 18 months after baseline.

Note: Impact estimates for subgroups are shown only if a test for differences in impacts among the subgroups met the study criterion for statistical significance ($p < .05$). For example, a test result indicated that the treatment effect on more than one lifetime partner (for sexual intercourse) was significantly different between the sexually experienced at baseline subgroups.

[a] This column shows the estimated treatment effect (treatment/control difference in percent responding affirmatively) for the subgroup indicated in the row.

[b] This column shows the statistical test result for whether the treatment effect for the subgroup indicated in the row was significantly different than zero.

* $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed tests).

Exhibit C.3 Short-Term Effects on Non-Behavioral Outcomes by Subgroup

	Treatment Effect [a]	*p*−Value [b]
Intention to use condom if they were to have sexual intercourse		
Subgroup: Respondent sexual experience at baseline		
Never sexually active at baseline (*n* = 125)	−6.77	.334
Ever sexually active at Baseline (*n* = 1,679)	7.50***	.000

Source: Follow-up survey administered nine months after baseline.

Note: Impact estimates for subgroups are shown only if a test for differences in impacts among the subgroups met the study criterion for statistical significance ($p < .05$). For example, a test result indicated that the treatment effect on knowledge of pregnancy risk was significantly different for younger versus older respondents.

[a] This column shows the estimated treatment effect (Treatment/control difference in the average percent of items answered correctly) for the subgroup indicated in the row.

[b] This column shows the statistical test result for whether the treatment effect for the subgroup indicated in the row was significantly different than zero.

* $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed tests).

Appendix D: Supporting Tables

Exhibit D.1: Characteristics of the Analytic Sample at Baseline (short-term follow-up)

Measure	Treatment Mean[a]	Control Mean	Group Difference[b]	*p*-Value
Demographic characteristics				
Age				
Mean	17.12	17.14	−0.01	.794
Race/ethnicity[c]				
Hispanic	16.80	19.90	−3.10	.086
Black	35.77	35.24	0.53	.807
White	33.58	31.65	1.93	.335
Other	13.85	13.21	0.64	.697
Family structure and relationships				
Lives with biological parents	78.97	78.02	0.96	.630
Feels very close to and cared for by father	29.58	26.07	3.51	.139
Feels very close to and cared for by mother	44.25	48.33	−4.09	.101
Risk behaviors				
Ever smoked cigarettes	51.49	53.28	−1.79	.465
Ever drank alcohol	78.54	82.10	−3.56	.071
Ever used marijuana	67.29	68.03	−0.74	.750
Sexual activity				
Currently sexually active (in last 90 days) [d]	82.97	83.53	−0.56	.761
Sexual intercourse in the last 90 days	78.70	79.08	−0.38	.852
Oral sex in the last 90 days	66.00	66.28	−0.28	.905
Anal sex in the last 90 days[d]	11.52	10.68	0.84	.637
Sexual risk behavior				
Sexual intercourse without a condom in the last 90 days	59.86	59.14	0.72	.767
Oral sex without a condom in the last 90 days	62.36	62.15	0.22	.927
Anal sex without a condom in the last 90 days[d]	9.58	7.91	1.68	.297
Sexual intercourse without birth control in the last 90 days	31.31	31.47	−0.15	.947
Sexual intercourse with more than one partner (lifetime)	66.21	67.35	−1.15	.626
Sexual intercourse with more than 5 partners (lifetime)	23.60	23.37	0.24	.911
Knowledge[e]				
Knowledge of pregnancy risk	68.50	70.96	−2.47	.173
Knowledge of STI risk	68.75	67.24	1.51	.246
Attitudes[f]				
Attitudes toward protection	3.26	3.26	0.00	.971
Intentions				
Intentions to have oral sex in the next 12 months	60.34	61.82	−1.47	.528
Intentions to have sexual intercourse in the next 12 months	84.22	84.40	−0.18	.917

Measure	Treatment Mean[a]	Control Mean	Group Difference[b]	*p*-Value
Intentions to use a condom if they were to have sexual intercourse in the next 12 months	84.62	83.91	0.72	.693
Intentions to use birth control if they were to have sexual intercourse in the next 12 months	92.76	91.48	1.29	.326

Source: Baseline survey administered prior to randomization.
Note: The baseline treatment-control difference was estimated where the dependent variable was the baseline measure, and the only independent variables included in the model were the treatment group indicator and terms for the randomization blocks. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups. Results in this table are based on the analytic sample of 1,786–1,809 respondents who provided valid short-term survey responses to relevant items except for the items measuring how close the respondent feels to their mother (n = 1,779) and father (n = 1,603), number of partners (n = 1,710), and anal sex (n = 1,379). Values shown are percentages unless otherwise indicated. The items that compose measures of attitudes toward risky sexual behavior, motivation to delay childbearing, refusal skills, and condom negotiation skills were not asked at baseline.
[a] The treatment mean was calculated as the sum of the control group mean and the model estimated treatment-control difference (group difference).
[b] *Other* is defined as Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, multiracial, or undisclosed race.
[c] Knowledge variables are composite scale scores representing the proportion of items answered correctly.
[d] Attitude variable is a composite scale score with higher scores indicating more positive attitudes.
[e] *Sexual activity* is defined differently across grantees. In Hennepin County and Planned Parenthood of Greater Orlando, sexual activity refers to sexual intercourse, oral sex, and anal sex. Youth were not asked about anal sex in Knox County.
* $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed tests).

Exhibit D.2: Characteristics of the Analytic Sample at Baseline (longer-term follow-up)

Measure	Treatment Mean[a]	Control Mean	Group Difference[b]	*p*-Value
Demographic characteristics				
Age				
Mean	17.14	17.17	−0.03	.567
Race/ethnicity[c]				
Hispanic	16.39	19.80	−3.42	.060
Black	35.62	34.53	1.08	.618
White	34.32	32.08	2.24	.267
Other	13.68	13.58	0.10	.953
Family structure and relationships				
Lives with biological parents	78.49	77.65	0.84	.674
Feels very close to and cared for by father	28.87	26.01	2.85	.225
Feels very close to and cared for by mother	44.35	48.17	−3.82	.125
Risk behaviors				
Ever smoked cigarettes	51.65	54.28	−2.63	.284
Ever drank alcohol	78.78	83.53	−4.74*	.015
Ever used marijuana	66.74	68.59	−1.85	.426
Sexual activity				
Ever sexually active[d]	93.13	94.54	−1.40	.236
Currently sexually active (in last 90 days)[d]	83.16	83.25	−0.09	.962
Sexual intercourse in the last 90 days	78.65	79.10	−0.45	.822
Oral sex in the last 90 days	65.92	65.89	0.03	.990
Anal sex in the last 90 days[d]	11.68	10.33	1.35	.450
Sexual risk behavior				
Sexual intercourse without a condom in the last 90 days	60.35	60.70	−0.34	.887
Oral sex without a condom in the last 90 days	62.22	61.90	0.33	.889
Anal sex without a condom in the last 90 days[d]	9.73	7.69	2.04	.211
Sexual intercourse without birth control in the last 90 days	31.42	32.84	−1.42	.540
Sexual intercourse with more than one partner (lifetime)	64.98	67.34	−2.36	.315
Sexual intercourse with more than 5 partners (lifetime)	22.98	23.18	−0.20	.922
Consequences of sexual risk behavior				
Ever pregnant or gotten someone pregnant (lifetime)	19.70	14.99	4.71*	.011
Diagnosed with STI in the last 12 months	12.17	14.19	−2.03	.212
Knowledge[e]				
Knowledge of pregnancy risk	68.66	70.62	−1.96	.279
Knowledge of STI risk	68.72	67.14	1.58	.232
Attitudes[f]				
Attitudes toward protection	3.25	3.26	−0.01	.744
Intentions				
Intentions to have oral sex in the next 12 months	59.90	61.56	−1.67	.477
Intentions to have sexual intercourse in the next 12 months	84.03	84.35	−0.32	.860
Intentions to use a condom if they were to have sexual intercourse in the next 12 months	84.27	83.20	1.07	.558
Intentions to use birth control if they were to have sexual intercourse in the next 12 months	92.49	91.25	1.24	.353

Source: Baseline survey administered prior to randomization.
Note: Results in this table are based on the analytic sample of 1,368 – 1,815 longer-term survey respondents who provided valid survey responses to relevant items on the baseline survey. Values shown are percentages unless otherwise indicated. The items that compose measures of attitudes toward risky sexual behavior, motivation to delay childbearing, refusal skills, and condom negotiation skills were not asked at baseline.
[a] The treatment mean was calculated as the sum of the control group mean and the model estimated treatment-control difference (group difference).
[b]The baseline treatment-control difference was estimated where the dependent variable was the baseline measure, and the only independent variables included in the model were the treatment group indicator and terms for the randomization blocks. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.
[c] Racial ethnic categories are Hispanic, Black non-Hispanic, White non-Hispanic, and Other race non-Hispanic, where Other is defined as Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, multiracial, or undisclosed race.
[d] *Sexual activity* is defined differently across grantees. In Hennepin County and Planned Parenthood of Greater Orlando, *sexual activity* refers to sexual intercourse, oral sex, and anal sex. Youth were not asked about anal sex in Knox County.
[e]Knowledge variables are composite scale scores representing the proportion of items answered correctly.
[f]Attitude variable is a composite scale score with higher scores indicating more positive attitudes.
* $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed tests).

Appendix E: Behavioral Impacts for Sexually Inexperienced Youth

Exhibit E.1: Short-Term Effects on Sexual Behavior and Sexual Risk for Youth Sexually Inexperienced at Baseline

Outcome	Adjusted Treatment Mean[a]	Unadjusted Control Mean	Treatment Effect[b]	*p*-Value
Sexual activity (percentage responding affirmatively) [c]				
Initiation of sexual activity	29.10	50.00	−20.90	.097
Currently sexually active (in last 90 days)	22.06	32.50	−10.44	.338
Sexual intercourse in the last 90 days	18.07	27.50	−9.43	.329
Oral sex in the last 90 days	11.69	22.50	−10.81	.235
Anal sex in the last 90 days	3.18	7.14	−3.97	.591
Sexual risk (percentage responding affirmatively)				
Sexual intercourse without birth control (in last 90 days)	5.51	10.00	−4.49	.506
Sexual intercourse without a condom (in last 90 days)	13.00	12.50	0.50	.950
Oral sex without a condom (in last 90 days)	12.79	20.00	−7.21	.402
Anal sex without a condom (in last 90 days)	2.38	3.57	−1.19	.777
Sexual intercourse with more than one partner (lifetime)	9.14	30.56	−21.42*	.028
Sexual intercourse with more than five partners (lifetime)	6.38	5.56	0.82	.869

Source: Follow-up survey administered nine months after baseline.
Note: Results in this table are based on 124 respondents who provided valid survey responses to relevant items, except for the items measuring number of partners (n = 115) and anal sex (n = 83).
[a] The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression adjusted impact estimate (treatment effect).
[b] The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. The treatment effect is expressed as a difference in percentage points. Due to rounding, reported treatment effects may differ from differences reported between reported means for the treatment and control groups.
[c] Sexual activity is defined differently across grantees. In Hennepin County and Planned Parenthood of Greater Orlando, sexual activity refers to sexual intercourse, oral sex, and/or anal sex. Youth were not asked about anal sex in Knox County.
* $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed tests).

Exhibit E.2: Longer-Term Effects on Sexual Activity and Sexual Risk Behavior for Youth Sexually Inexperienced at Baseline

Outcome	Adjusted Treatment Meana	Unadjusted Control Mean	Treatment Effect[b]	*p*-Value
Sexual behavior				
Sexual activity (percentage responding affirmatively) [c]				
Initiation of sexual activity	57.95	45.95	12.01	.348
Currently sexually active (in last 90 days)	45.49	32.43	13.06	.288
Sexual intercourse in the last 90 days	34.86	18.92	15.94	.141
Oral sex in the last 90 days	25.42	21.62	3.80	.715
Anal sex in the last 90 days	5.53	8.70	−3.16	.711
Sexual risk (percentage responding affirmatively)				
Sexual intercourse without birth control (in last 90 days)	7.65	8.11	−0.46	.922
Sexual intercourse without a condom (in last 90 days)	15.25	18.92	−3.67	.706
Oral sex without a condom (in last 90 days)	21.50	21.62	−0.12	.991
Anal sex without a condom (in last 90 days)	6.37	8.70	−2.32	.753
Sexual intercourse with more than one partner (lifetime)	38.98	21.62	17.36	.125
Sexual intercourse with more than five partners (lifetime)	11.23	2.70	8.53	.171
Sexual consequences (percentage responding affirmatively)				
Pregnant since baseline	11.47	5.41	6.06	.438
Diagnosed with STI in the last 12 months	6.74	0.00	6.74	.225

Source: Follow-up survey administered 18 months after baseline.
Note: Results in this table are based on 122–123 respondents who provided valid survey responses to relevant items, except for the items measuring anal sex (n = 80).
[a] The treatment group mean is regression-adjusted, calculated as the sum of the unadjusted control group mean and the regression adjusted impact estimate (treatment effect).
[b] The treatment effect was estimated in a one-level fixed-effects regression model that controls for randomization blocks and other covariates. The treatment effect is expressed as a difference in percentage points.
[c] Sexual activity is defined differently across grantees. In Hennepin County and Planned Parenthood of Greater Orlando, sexual activity refers to sexual intercourse, oral sex, and anal sex. Youth were not asked about anal sex in Knox County.
* $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed tests).