

Teen Pregnancy Prevention Replication Study: Short-Term Impacts of *Safer Sex Intervention*

|EVALUATION REPORT|

OCTOBER 2016

Safer Sex Intervention: Short-Term Impact Report

Teen Pregnancy Prevention Replication Study

October 2016

Prepared for:
Amy Farb
The Office of Adolescent Health

Lisa Trivits
Sarah Oberlander
The Office of the Assistant Secretary for Planning and Evaluation

Submitted by:
Meredith Kelsey(Abt)
Jean Layzer(Belmont Research Associates)
Cristofer Price(Abt)
Randall Juras (Abt)
Abt Associates
55 Wheeler St.
Cambridge, MA 02138

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

Recommended Citation

Kelsey, M., Layzer, J., Price, C., and Juras, R. “Safer Sex Intervention Short-Term Impact Report: Fndings from the Teen Pregnancy Prevention Replication Study”. Cambridge, MA: Abt Associates Inc., October 2016.

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 provided guidance and support during the initial study design phase, and Fatih Unlu provided valuable feedback on the report. Thanks to Dan Weiss and Heather He for excellent programming assistance, and Stefanie Falzone and Jan Nicholson for producing 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) are a priority for the U.S. Department of Health and Human Services (DHHS). 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) within DHHS, 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 Teen Pregnancy Prevention (TPP) Program was authorized in 2010 as part of the larger Teen Pregnancy Prevention Initiative. The program initially included $100 million in funds annually to support programming. Of these funds, $75 million were available 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 teen pregnancy, STIs, and sexual risk behaviors (Kappeler and Farb, 2014).

The TPP Program also 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). Evidence for many of the programs identified in this review is based on a single study 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 on the ground. For these reasons, a carefully designed study of multiple replications of different program models is an important contribution to the existing research.

1.1 The Replication Study

The TPP Replication Study[1] is being conducted for OAH, under a contract with the Office of the Assistant Secretary for Planning and Evaluation (ASPE), 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 study tests multiple replications of three popular evidence-based program models to determine their effectiveness across different settings and populations. The strategy of selecting multiple replications of each program model increases the generalizability of the findings. It also allows us to pool data across the three replication sites to assess impacts on such behavioral outcomes as pregnancy and for key subgroups (e.g., age and sexual experience). In addition, the strategy lets us examine variation in impacts across replications for each program model and provides evidence about the generalizability of program effectiveness.

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

Implementation Study. A comprehensive Implementation Study will provide information about the contexts in which the evidence-based programs were implemented and the challenges faced in implementing them. It will also allow us to assess aspects of program implementation that are associated with program impacts.

1.2 The Three Models Replicated

OAH, with its ASPE partners, selected three program models from the first round of TPP-funded grants to test and replicate: the *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 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, which focuses on the *Safer Sex Intervention (SSI),* is one in a series of reports that present findings on the implementation and effectiveness of the three program models. The report presents findings from the first of two follow-up surveys designed to examine the short-term and longer-term impacts of *SSI.* Two companion reports examine the short-term impacts (6 to 12 months post-baseline[3]) of the other program models in the study (*Reducing the Risk* and *¡Cuídate!*). Three final impact reports will present findings on the longer-term impacts (18 to 24 months post-baseline) of all three program models. Three Implementation Study reports will document the implementation of each of the 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.[4]

---

[2] The Teen Outreach Program (TOP) is 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] Where "baseline" means the point at which a study participant entered the study.

[4] See https://aspe.hhs.gov/basic-report/tpp-replication-study

2. The Program Model: *Safer Sex Intervention (SSI)*

The *Safer Sex Intervention (SSI)* is a clinic-based intervention intended to reduce the incidence of STIs and increase condom use among high-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 (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). In addition to the first 50-60 minute session, three subsequent booster sessions, similar in content, are delivered one, three and six months after the initial session. These sessions can vary in length from 10-20 minutes, depending on the needs and interest of the client, and are used to review information, assess progress and provide additional information and practice, if needed. Participants are offered condoms and informational materials.

Exhibit 2.1: Wheel of Change

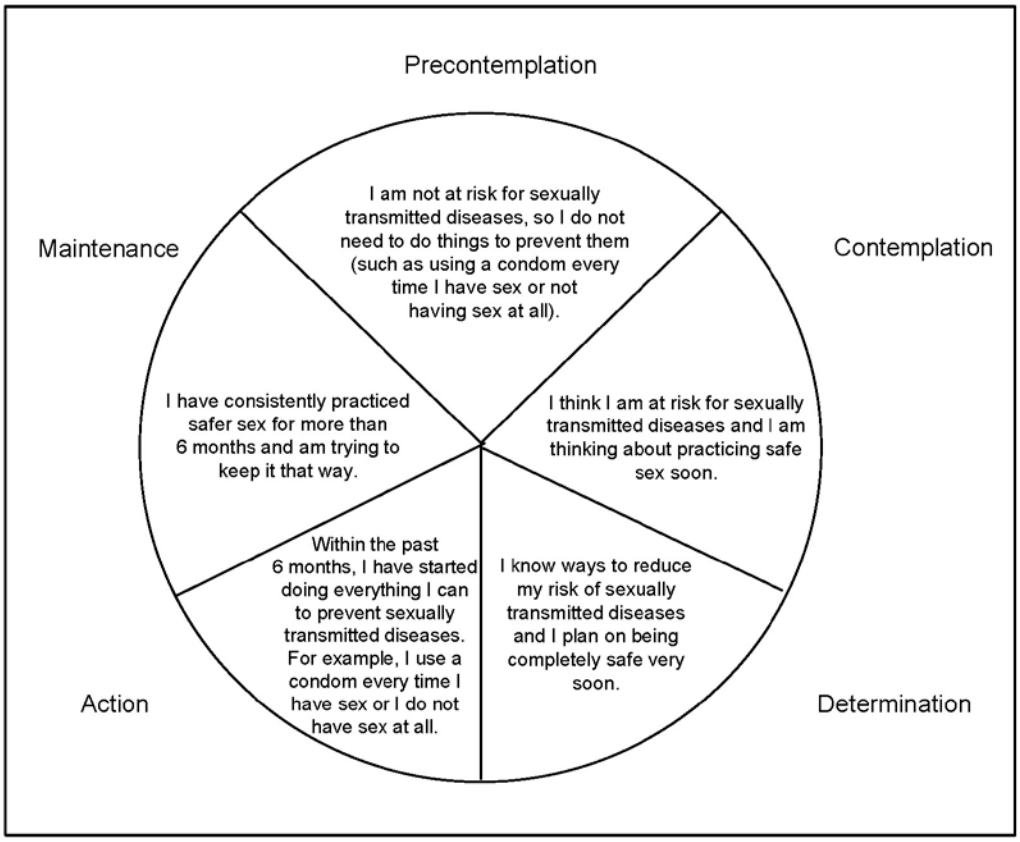


Exhibit 2.2: *SSI*: Initial Session, Topics and Core Elements Covered

Initial Session	Topic/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, R., Rex, P., & Shrier, L. (2011). *Safer Sex Intervention* adaptation kit. Scotts Valley, CA: ETR Associates.

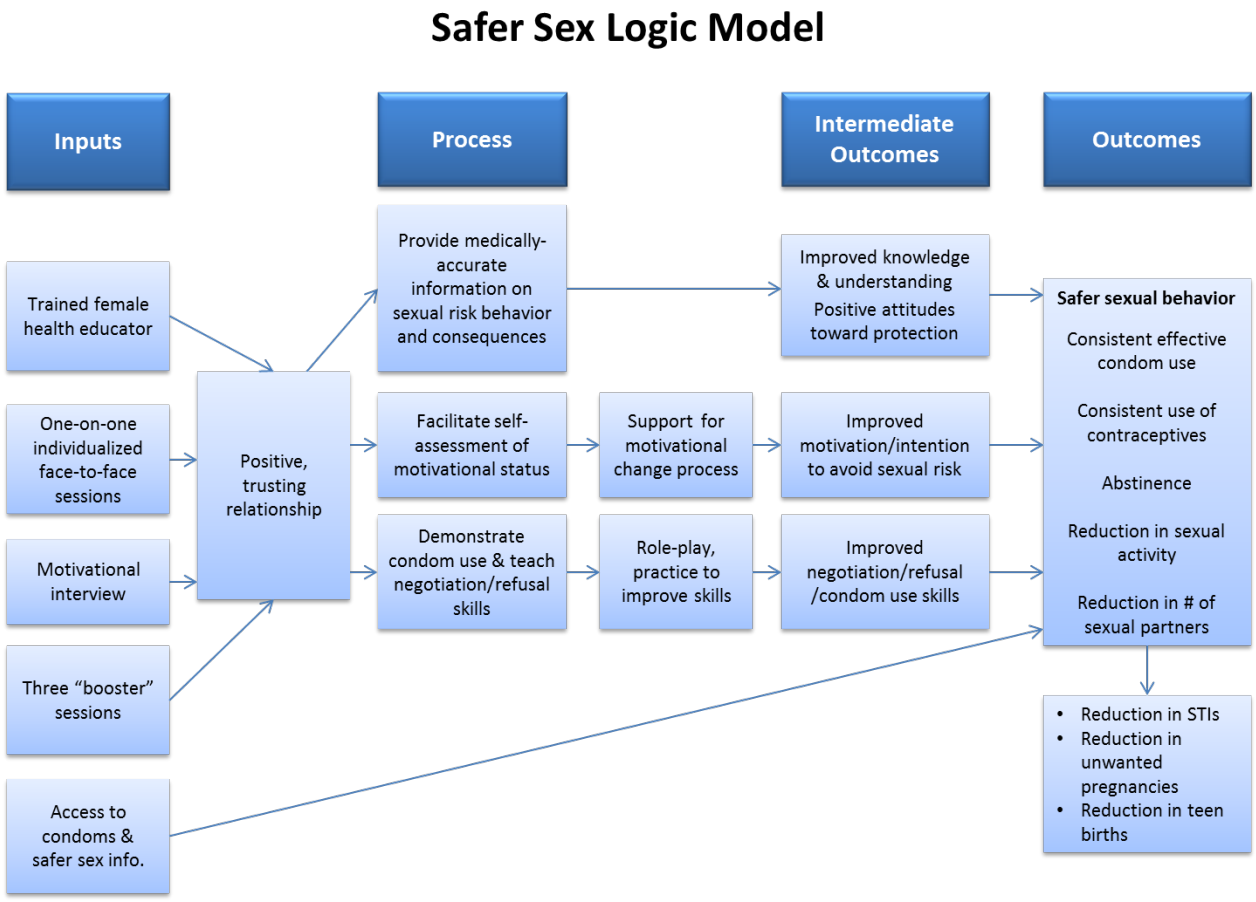
2.1 *Safer Sex Intervention* Logic Model

The theoretical framework for *SSI* draws on Social Cognitive Theory, the Transtheoretical Model of Behavior Change, and the technique of motivational interviewing. 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 rare in this field, although it is widely used in other fields and with adolescents. 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. 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.

Exhibit 2.3 shows the program elements, the 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: namely, 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, pregnancies, and births among teens.

Exhibit 2.3: Logic Model for *SSI*



In Section 3.2.2, we describe in more detail the modifications to the program model proposed, approved, and implemented by the three grantees that implemented the program replications. Modifications were made to update materials and improve 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 (DHHS, 2010). As with many other program models identified through this review, evidence for *SSI*'s effectiveness comes from a single study (Shrier et al., 2001) that was completed many years ago.

The program was originally 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 non-program participants 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

The impact study is designed to estimate the effects of three replications of *SSI*.[5] It addresses questions about the effects of the program on participants' sexual risk behaviors, as well as on the intermediate outcomes the logic model predicts will lead to the behavioral outcomes that *SSI* seeks to achieve.

The current report focuses on short-term program effects nine months after study enrollment. It is guided by the research questions below.[6]

3.1 Research Questions

- *Did SSI increase teens' exposure to information on reproductive health, contraceptive methods, and STI transmission and prevention?*
- *Did SSI improve teens' knowledge of pregnancy risk and STI risk?*
- *Did SSI have positive effects on teens' attitudes toward protection (i.e., birth control and condom use) or risky sexual behavior?*
- *Did SSI increase motivation to delay childbearing?*
- *Did SSI decrease intentions to engage in risky sexual behavior?*
- *Did SSI increase teens' confidence in their ability to refuse unwanted sex (refusal skills) or to negotiate safe sex (condom negotiation skills)?*
- *Did SSI reduce sexual behavior and sexual risk?*
- *Do program impacts differ by replication site and for key subgroups (e.g., age, race/ethnicity, sexual experience at baseline)?*

3.2 Study Design

In each of the replication sites, the study employed an experimental design in which young women were randomly assigned to a group that was offered *SSI* or to a control group that received the clinic's usual standard of care. This section describes the selection of the three replication grantees, site-specific program designs, settings for the program, recruitment and random assignment, and the treatment and control conditions.

3.2.1 Selection of Replication Grantees

The study design called for evaluating at least three replications of the model. At the time of site selection for the study, *SSI* was being replicated by at least five grantees. Most of these grantees had not planned

[5] A more detailed impact study design report can be found at https://aspe.hhs.gov/basic-report/tpp-replication-study.

[6] The final impact report will answer a similar set of questions about program effects on risk behaviors after 18 months. It will also examine program effects on pregnancy.

for a rigorous evaluation, a fact that complicated grantee recruitment selection into the study.[7] One of the five grantees was eliminated due to its low service volume and 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 were:

- **Hennepin County Human Services and Public Health Department.** Hennepin County Human Services and Public Health Department (Hennepin County) has played a longstanding 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 the prevention of 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.** Knox County Health Department (Knox County) is the local public health agency serving the City of Knoxville and Knox County. The Department's 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.** Planned Parenthood of Greater Orlando (PPGO), an affiliate of Planned Parenthood Federation of America, Inc., operates as a community based non-profit 501(c)(3) organization.[8] Since 1995, the organization 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.

3.2.2 Site-Specific Program Designs

In all three replication sites, *SSI* grantees served sexually-active young women or young women considering sexual initiation. 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

---

[7] 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 were not expected to have a rigorous local evaluation. Hennepin County, a larger-scale replication, had proposed a rigorous local evaluation.

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

the original study).[9] 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 the developer approved this adaptation for all three *SSI* grantees.[10] Other approved adaptations that were implemented in all three replication sites included replacing the original video, which was outdated, as recommended by the developer.[11] Two of the three replication sites (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 their 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*.

3.2.3 Settings for the Program

SSI was implemented in clinics within each of the replication sites. In each of the clinics, the intervention was seen as 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, the largest replication site, offered *SSI* in 19 different clinics during the study enrollment period. The County contracted with provider agencies to deliver *SSI* to at-risk youth from different 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 homeless youth. Clinics were located throughout the county in eight cities. The clinics varied in geographic location (urban vs. suburban) and the populations served. Each clinic had individual targets for recruitment based on their number of full-time employees (FTEs).

---

9 Hennepin County had originally proposed serving males as well as females, but this adaptation was not approved by OAH.

10 In all three replication sites, the grantees worked with the developer during the grant proposal phase. Upon award, OAH recommended that the developer be involved, and each of the replication sites established consulting agreements with her. 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 (FAQs) she received from grantees prior to the availability of the adaptation kit.

11 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: It should be brief, include peers, demonstrate correct condom use and preferably use humor or otherwise be entertaining (*Safer Sex Intervention* Frequently Asked Questions: correspondence from Lydia Shrier, September 22, 2011). A more detailed description of the videos and adaptations is included in the *SSI* Implementation Report (forthcoming).

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 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 counties. The populations served by the clinics are predominantly White Non-Hispanic; only one clinic serves a significant proportion of Hispanic/Latino families.

Planned 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.

3.2.4 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 if 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.[12] 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.

[12] 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 was 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.

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.[13] 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 and next steps.

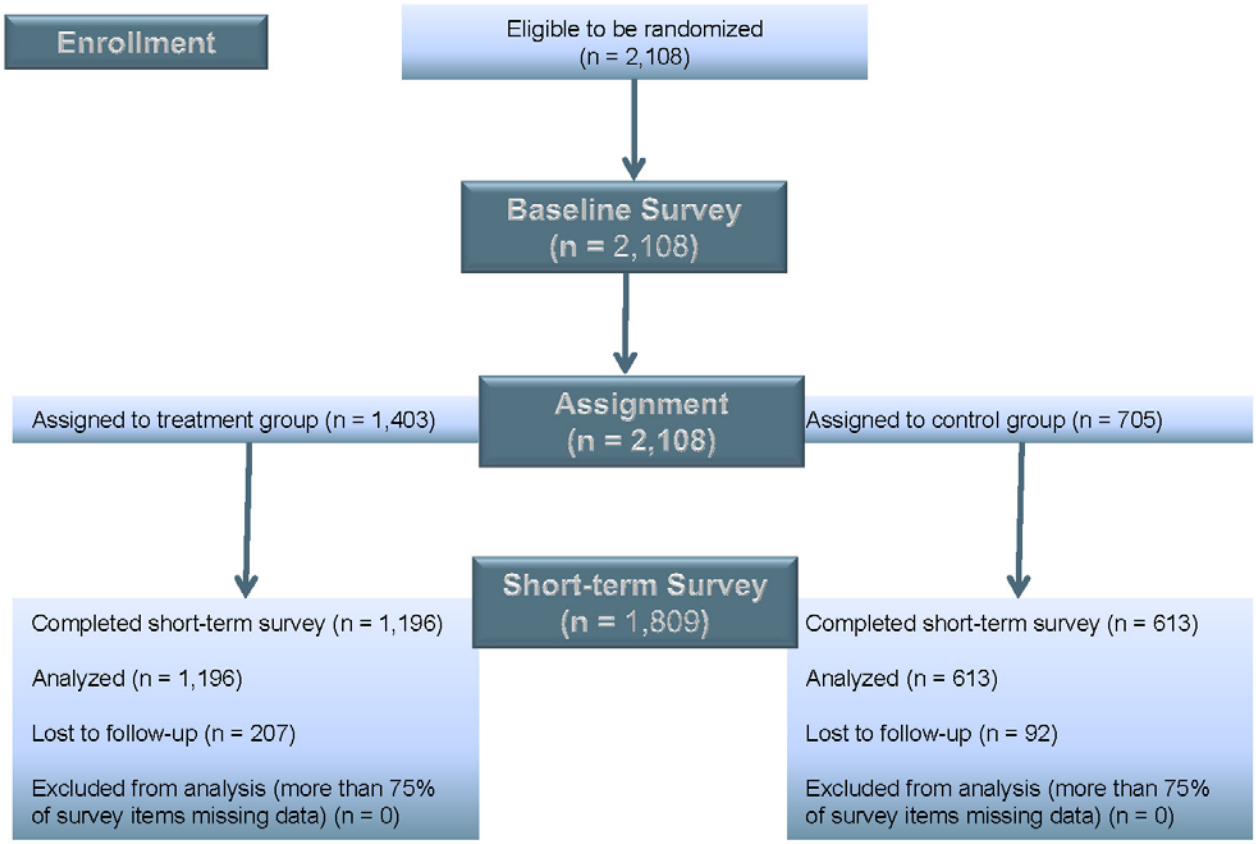
Random assignment occurred independently in each of the clinic sites. Individual sample members within clinics were randomly assigned on a rolling basis within randomization blocks based on site and age (less than 15 years old versus 15 years or older) and time (3-6 month periods). The randomization 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 designed to check for duplicate participants across the 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 3.1 shows, across the three replication sites, 2,108 young women were eligible for and consented to the study.[14] All who consented (100 percent) completed a baseline survey and were then randomized, and 1,809 (85.8 percent) completed the short-term follow-up survey.

[13] 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.

[14] 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.

Exhibit 3.1: Study Sample



3.2.5 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 their usual standard of care (Exhibit 3.2).

Exhibit 3.2: 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	Individualized sessions with trained health educator; initial session and boosters at 1-, 3-, and 6-month intervals	Standard of care
Planned Parenthood of Greater Orlando 2 clinics in Orlando, FL	Individualized sessions with trained health educator; initial session and boosters at 1-, 3-, and 6-month intervals	Standard services for those recruited in the clinic; pregnancy or STI test for those recruited outside the clinic

3.3 Measures and Data Collection Strategies

3.3.1 Data Collection Strategy

To assess the impacts of the intervention, young women in each of the three replication sites were surveyed three times: at baseline, before the intervention began; nine months after the baseline survey (short-term follow-up); and 18 months after the baseline survey (longer-term follow-up). Each time, 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 the oversaw baseline survey and provided gift cards afterward. As Exhibit 3.3 shows, all 2,108 study participants completed a baseline survey.

For the short-term follow-up survey (9 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 e-mail and text messages before the survey went live and throughout the survey period.[15] 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 on-line or helped them to access the survey. Gift cards were mailed to participants after completion.

As Exhibit 3.1 shows, a large majority (85.8 percent) of these young women subsequently completed the short-term survey. There was almost no difference in the response rates of youth in the treatment group versus those in the control group. Of the three sites, PPGO had the highest response rates.

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

Exhibit 3.3: *SSI* Short-term Survey Response Rate

	Total N	Eligible Participants		First Follow-up Completes Total		First Follow-up Completes Treatment		First Follow-up Completes Control	
		Treatment	Control	N	Percent	N	Percent	N	Percent
All Sites	2108	1403	705	1809	85.8%	1196	85.3%	613	87.0%
Hennepin County	1177	785	392	968	82.2%	639	81.4%	329	83.9%
Knox County	491	326	165	413	84.1%	275	84.4%	138	83.6%
Planned Parenthood of Greater Orlando	440	292	148	428	97.3%	282	96.6%	146	98.7%

3.3.2 Measures

The first follow-up survey collected information from youth on a variety of factors, including questions that allow us to measure three sets of outcomes: (1) exposure to information about topics related to sexual risk behavior; (2) intermediate outcomes (i.e., factors that are believed to lead to behavioral outcomes); and (3) sexual behavior and sexual risk. We briefly describe these measures here. A more complete description of these measures and the individual survey items they comprise can be found in Appendix D. Exhibit 3.4 summarizes the outcome measures and their construction.

Exposure to Sexual Health Information. In the first follow-up survey, we asked youth about their exposure to information about reproductive health and related topics. Youth were asked if they had received information about any of a set of topics in the 12 months preceding the survey. Because the topics were distinct, we examined responses to individual survey questions rather than creating and analyzing a composite measure.

Intermediate Outcomes. Drawing on knowledge of the program's theory of change and exploratory factor analysis, we constructed composite measures to assess four factors that potentially lead to behavioral outcomes: (1) ***knowledge*** of pregnancy risk and STI risk; (2) ***attitudes*** toward protection (i.e., use of condoms and other birth control methods) and attitudes toward risky sexual behavior; (3) ***motivation*** to delay childbearing; and (4) refusal and condom negotiation ***skills***. For a fifth measure, we analyzed four single-item measures: (5) ***intentions*** to engage in risky sexual behavior.

Knowledge. We constructed two composite measures: *knowledge of pregnancy risk* and *knowledge of STI risk.* The four items that make up the first measure and the twelve items that make up the second are all factual questions, testing youth's knowledge of the circumstances under which a woman can become pregnant (e.g., "A woman is protected from pregnancy the day she begins taking the pill") and the effectiveness of condoms and other methods of birth control in preventing pregnancy (e.g., "If birth control pills are used correctly and consistently, how much can they decrease the risk of pregnancy?"), as well as facts about STIs and their transmission (e.g., "You can't get infected with HIV if you have sex only once or twice without a condom"). All items were scored 1 for a correct answer and 0 for an incorrect answer; scores were averaged across the items that make up a measure and multiplied by 100 to indicate the percentage of items answered correctly.

Attitudes. We constructed two composite measures of attitudes: *attitudes toward protection* and *attitudes toward risky sexual behavior.* For the twelve items that make up the first measure (attitudes toward protection), youth were asked if they agreed or disagreed with statements such as "birth

control is important to make sex safer." Four response categories ranging from "strongly disagree" to "strongly agree" were scored from 1 to 4 and then scores for individual items were averaged, with higher values representing more positive attitudes toward the use of protection. For the seven items that comprise the second measure (attitudes toward risky sexual behavior), youth were asked if they agreed or disagreed with statements such as "It's OK to have sex with someone on the first night you meet them." Responses were scored 0 (disagree) or 1 (agree) and averaged across the items and multiplied by 100 to indicate the percentage of items agreed with, with higher scores representing higher levels of support for risky behavior.

Motivation. We constructed one composite measure, *motivation to delay childbearing*, which includes three survey items that asked respondents if they agreed or disagreed with statements such as "It is important for you to finish school before you have a child." Four response categories ranging from "strongly disagree" to "strongly agree" were scored from 1 to 4. Scores for individual items were averaged, with higher values representing greater motivation to delay childbearing.

Skills. We constructed two composite measures of skills: *refusal skills* and *condom negotiation skills*. The measure of *refusal skills* comprised six items probing respondents' perceptions of their ability to refuse to engage in sexually risky behavior (e.g., "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?"). Possible responses ranged from "I'm sure I could not" to "I'm sure I could" and were coded 1 to 4. Scores for individual items were averaged, with higher scores representing greater certainty of refusal skills. The measure of *condom negotiation skills* included seven items asking about respondents' perceptions of their ability to obtain and negotiate the use of condoms with a partner (e.g., "If you were going to have sex, could you insist on using a condom even if your partner didn't want to use one?"). Possible responses ranged from "I'm sure I could not" to "I'm sure I could," coded 1 to 4. All items were averaged, with higher scores representing greater certainty of condom negotiation skills.

Intentions. We included four single item measures of youth's intentions related to sexual activity in the year following the survey. The first item asked about oral sex; the second about sexual intercourse; the third about condom use; and the fourth about birth control. Responses to each of the four items were scored 0 or 1 (1 for those responding that their intentions were "probably" or "definitely").

Short-term Behavioral Outcomes. To address the study's most important questions about the impact of the intervention, we identified 10 measures in the domain of youth sexual behavior at the short-term follow-up: (1) *currently sexually active (in the last 90 days);* (2) *sexual intercourse in the last 90 days;* (3) *oral sex in the last 90 days;* (4) *anal sex in the last 90 days;* (5) *sexual intercourse without birth control (in last 90 days);* (6) *sexual intercourse without a condom (in last 90 days);* (7) *oral sex without a condom (in last 90 days);* (8) *anal sex without a condom (in last 90 days);* (9) *sexual intercourse with more than one partner (lifetime);* and (10) *sexual intercourse with more than five partners (lifetime).*

The first measure of sexual activity is defined differently across replication sites. In Hennepin County and PPGO, "sexual activity" refers to sexual intercourse, oral sex, and/or anal sex. Youth were not asked about anal sex in Knox County; in this site sexual activity refers to sexual intercourse and/or oral sex. The first eight measures are single items, with yes/no answers. The last two, *sexual intercourse with more than one partner* and *sexual intercourse with more than five partners*, reflect responses to a question about the total number of people with whom youth had ever engaged in sexual intercourse. *Sexual*

intercourse with more than one partner 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* was coded 0 or 1, with 1 representing six or more sexual partners and 0 representing five or fewer sexual partners in one's lifetime.

Exhibit 3.4: Outcome Measures

Measure	Definition
EXPOSURE TO INFORMATION & INTERMEDIATE OUTCOMES	
Domain: Exposure to Information	
Exposure to Sexual Health Information	Eight single items reflecting exposure to information about: (a) relationships or marriage; (b) abstinence from sex; (c) birth control methods; (d) where to obtain birth control; (e) STIs; (f) how to talk with a partner about sex and birth control; (g) how to say no to sex; and (h) how babies are made. Responses were coded as 1="yes" and 0="no."
Domain: 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. Average scores, multiplied by 100, 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. Scores range from 0 to 100 and represent the percent of the 12 questions answered correctly, with higher values representing more accurate knowledge.
Domain: Attitudes	
Attitudes toward protection	Continuous index: average of responses to 12 questions about attitudes towards 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.
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 percent of items agreed with). Scores range from 0 to100 with higher values representing more support for risky behavior.
Domain: 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.
Domain: Intentions	
Intentions to have oral sex in the next year	Single item scored 0 or 1, with 1 representing stronger intention
Intentions to have sexual intercourse in the next year	Single item scored 0 or 1, with 1 representing stronger intention
Intentions to use a condom if they were to have sexual intercourse in the next 12 months	Single item scored 0 or 1, with 1 representing stronger intention
Intentions to use birth control if they were to have sexual intercourse in the next 12 months	Single item scored 0 or 1, with 1 representing stronger intention

Measure	Definition
Domain: 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
Domain: Youth Sexual Behavior and Sexual Risk at the Short-Term Follow-Up	
Currently sexually active (in last 90 days)* 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 intercourse without birth control (in last 90 days) * 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

*Designated as a confirmatory outcome, as discussed in Section 3.4.5

3.4 Analytic Approach

The impact analysis examines the extent to which *SSI* affected each of the study's outcomes. In testing for these effects, we use two-tailed hypothesis test procedures, because we do not want to rule out the possibility that the intervention might adversely affect one or more of the outcomes.

Our basic strategy for estimating program impacts is to compare the outcomes of treatment and control group members using a regression framework, in which we include 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 et al., 2009).

3.4.1 Estimation of Impacts for the Full Sample

In this section, we report impact estimates that are pooled across the three *SSI* replication sites. OAH's requirements to define, measure, and adhere to fidelity to the program model means that each of the three replication sites implemented the same core program elements. The random assignment and data collection procedures were also the same across all sites. These design elements ensure that impact estimates pooled at the program level represent rigorous tests of a well-defined and consistently implemented program model.

In this evaluation of *SSI*, 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 (3 to 6 month periods) in an approximate 2:1 treatment-to-control ratio. In order to account for the unequal ratio of treatment to control group members within blocks, and to ensure the impacts were estimated by comparing treatment group members to control group members within site, clinic, age, and time, block

dummy variables were included in the analytic model to represent the randomization blocks. This forces the analytic model to estimate an overall average impact that is a precision weighted average of estimated treatment effects within the site, clinic, age and time blocks. For each outcome, we estimate a model that reflects this design and has the basic structure of Equation 1.[16]

(1) $$Y_i = \beta_0 + \beta_1 T_i + \sum_{k=2}^{K+1} \beta_k X_{ki} + \sum_{m=1}^{M} \gamma_{0m} D_m + \varepsilon_i$$

In this model:[17]

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 a dummy 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 and STI risk knowledge, baseline intentions to have oral sex and sexual intercourse, and the baseline measure of the outcome when available.

D_m is the dummy variable representing the m^{th} randomization block. These block indicators reflect the fact that there were different treatment probabilities across blocks. Because random assignment blocks were constructed based on site, clinic, age, and time, the dummy variable also accounts for these factors.

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 coefficients on the covariates, β_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.[18]

Equation 1 estimates the impact of access to *SSI*. Because of the random assignment design, the systematic difference between the treatment and control groups is ***access*** to *SSI* services: Individuals

[16] 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).

[17] The analyses presented in this report used linear probability models for binary outcomes. 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 et al., 2010). There were no substantive differences in the inferences that results from any of the three modeling approaches.

[18] 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, 2008a, p. 70).

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, while control group members had access to the standard services available in the clinics and other services in the community. (See Section 3.2.5 for further details on the treatment and control conditions.) In the evaluation literature, the estimate of the average impact of access is referred to as the intent-to-treat (ITT) impact parameter. It measures the average impact on treatment group members 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, the ITT impact estimate will be very close to the impact on the members who actually participated in services.

Attendance data show that across all three sites, young women in *SSI* participated in enough of the program to meet the requirements imposed by OAH and the developer. Nearly all of the participants received the first session. In Knox County and PPGO, 99 percent of participants met with a health educator for the first session. In Hennepin County, scheduling constraints in some of the school-based clinics made it difficult to complete the baseline and the first session in the same day, and as a result, a small number of participants completed the baseline but weren't able to either stay or come back for the first session. Still, overall, a large majority in Hennepin County (89 percent) received the first session.

Finally, we report impact findings in tables showing the regression-adjusted treatment group mean, the unadjusted control group mean, and the difference between these two as the inferred regression-adjusted impact (and the p-value for the difference). For binary outcomes (e.g., condom use), we report impacts as percentage-point differences between the treatment and control group means. For all other outcomes, we show impact estimates in their original metric and additionally convert impact estimates to standardized effect sizes by dividing the impact estimate by the pooled standard deviation of the treatment and control groups, and we report these in a separate column.

3.4.2 Site-level Analyses

In addition to estimating impacts pooled across the three replication sites, we estimated impacts for each site separately and tested for differences in impacts across the three sites. We implemented these analyses by including treatment by site interaction terms in the model (i.e., Equation 1) and testing for the joint significance of the interaction terms.[19] When statistically significant differences in impact are found across sites for one or more outcomes, we discuss these differences in the main text of the report.[20] Site-specific impact estimates for all outcomes are presented in Appendix A.

---

[19] For the treatment-by-site interaction, a two degrees-of-freedom F test was used.

[20] The purpose of testing for differences across sites before discussing results in the main text is to guard against over-interpretation of spurious findings, some of which would be expected by chance in such a large group of outcomes. The basic idea behind the strategy of discussing site-specific impacts only when differences are found is that it is only credible to report an impact in one site—but not in another—if there is a significant difference between the sites. The site-specific results in Appendix A are not adjusted for multiple comparisons and any statistically significant findings should therefore be interpreted with caution.

3.4.3 Subgroup Analyses

In addition to the overall pooled impacts and site-level impacts, we estimated impacts for key subgroups of participants (based on age, race/ethnicity, and sexual experience 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) and testing for significance of the interaction term.[21]

To guard against potential over-interpretation of results among the very large number of subgroup estimates, we only present impact estimates for individual subgroups when there is a statistically significant difference between subgroups; e.g., the impact would only be presented for the subgroup of Hispanics if there were a statistically significant difference in impacts across racial/ethnic groups (see Appendix B).

To understand the effects of *SSI* among youth who were not sexually experienced at baseline, we examined the impacts on sexual behavior and sexual risk among this subgroup alone. These impacts were estimated using the same regression model used in the main impact analyses (see Equation 1) on the subset of youth who indicated they had never engaged in sexual intercourse, oral sex, and/or anal sex at baseline. The results for this small subgroup are presented in Appendix C.

3.4.4 Handling Missing Data

We used monetary incentives (gift cards) and intensive tracking to achieve the maximum possible response rate for both treatment and control groups, and have achieved very high response rates in each of the replication sites (see Exhibit 3.3).

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 indictors (or dummy variables) for such values were added to the impact analysis model (Puma et al., 2009).

3.4.5 Addressing Multiple Comparisons

Ongoing developments in the statistical analysis of the results of randomized trials emphasize that conventional statistical tests and confidence intervals apply to a single outcome. When analysts look over multiple outcomes for any statistically significant finding, the appropriate critical t-values are much higher; i.e., effects that appear to be statistically different from zero are not truly different from zero. In the literature, this is known as the problem of "multiple comparisons." Current guidance on how to approach this multiple comparison problem recommends distinguishing two categories of analyses (Schochet, 2008b). One—called "confirmatory tests"—includes a small number of critical outcome domains for which it is important to adjust error probabilities for multiplicity. Confirmatory analysis uses a high standard of evidence for deciding if an intervention has had its intended effect, in order for its findings to be considered conclusive rather than merely suggestive. A second category includes "exploratory tests" for which there is generally higher tolerance of errors and for which multiplicity adjustments may or may not be made.

---

[21] For the treatment-by-race/ethnicity interaction, a three degrees-of-freedom F test was used.

For this report, the impact analysis team pre-specified a multiple comparisons strategy that spans the two reports (i.e., the short-term and longer-term reports) and includes confirmatory and exploratory analyses. The confirmatory analysis seeks convincing evidence that *SSI* improved participants' behavioral outcomes past the end of the program. Before analyzing data, the research team pre-specified a small number of outcomes in three "domains," or sets of similar constructs, as part of the overall analytic strategy for both the short-term and longer-term reports. The three confirmatory outcome domains are: *recent sexual behavior at the short-term follow up*, *recent sexual behavior at the long-term follow up*, and *pregnancy*.

To control for multiple comparisons within each of the confirmatory domains, we applied a formal multiple comparisons correction (in particular, a Benjamini-Hochberg correction, as described in Appendix G of the What Works Clearinghouse Procedures and Standards Handbook, version 3.0, which which controls for the false positive rate by adjusting *P*-value thresholds). The correction does not affect the p-values that appear in tables of results, but it does change the interpretation of statistical significance. In particular, it raises the bar for rejecting the null hypothesis.

Two outcomes in the short-term report, **currently sexually active** (defined as sexual intercourse, oral sex, or anal sex in the last 90 days) and **sexual intercourse without birth control** (in last 90 days), were pre-specified as key outcomes in one of the study's three confirmatory outcome domains, *youth sexual behavior at the short-term follow up*. The other two domains, *recent sexual behavior at the long-term follow up* and *pregnancy* will be analyzed in the longer-term report, along with the findings presented here.

The exploratory analysis encompasses all other outcomes and research interests in the short-term report, e.g. impacts on intermediate outcomes and impacts on other behavioral outcomes. Given the large number of hypothesis tests that constitute the exploratory analysis, some false positive findings are to be expected. We do not make formal adjustments for multiple comparisons when reporting on statistical significance. However, to aid in interpretation, we specify the number of tests that were conducted (within domains) and the number of false rejections that would be expected given the number of tests if there were no impacts of treatment.

4. Results

This study is designed to determine whether *SSI* helps young women develop the knowledge, attitudes and skills to act in ways that ultimately protect them from the consequences of sexual risk behavior, such as STIs and unintended pregnancy. The program, when delivered with fidelity, is intended to provide information and affect potential intermediate outcomes such as: knowledge and understanding of reproductive health and avoidance of sexual risk; attitudes toward using protection; motivation to delay pregnancy; intentions to be sexually active and use protection; and skills needed to avoid sexual risk. The ultimate goals are reduced rates of unprotected sexual activity and unplanned pregnancy.

The short-term findings (nine months post-baseline) discussed here suggest that, across the three sites:

- The *SSI* program was implemented as intended.
- It was effective in changing some intermediate outcomes: attitudes toward protection, intentions to use condoms and perceived refusal skills.
- It was also effective in improving a primary behavioral outcome: use of birth control during sexual intercourse. There were no other improvements in reported sexual behavior or sexual risk behaviors.
- There were site-level differences, with a positive program impact on certain risk behaviors in one site; there were positive program impacts on certain risk behaviors for subgroups based on age, race/ethnicity, and sexual experience at baseline.

In this section, we expand on our conclusion that the intervention was indeed implemented with fidelity across replication sites, and then discuss findings for the full study sample and for individual sites, as well as any important findings for specific subgroups of young women (i.e., age, race/ethnicity, and sexual experience at baseline).

In addition to the exhibits in this section, tables documenting the site-level analyses can be found in Appendix A, and the corresponding tables documenting subgroup analyses can be found Appendix B.

4.1 Program Implementation

As we noted at the beginning of this report, a separate 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.

What we have learned from the Implementation Study that is directly relevant for this short-term impact report is that the intervention 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 the core elements and without modifications that threatened those core elements; and attendance was generally strong.

4.1.1 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.[22] 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 felt aligned with the philosophies of motivational interviewing and who understood that the health educator role was more about listening and eliciting conversation than about strictly educating. 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 and, in all three sites, program managers developed additional training specifically to supplement the intervention materials on 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.

4.1.2 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 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 supervisory staff 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—this was often 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, it is true that each grantee 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 were perceived to be the reasons for missed sessions. PPGO identified transportation as a substantial barrier to participation, so they 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

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

approval for remote video for the booster sessions, but they were less successful in implementing this adaptation. Knox County extended clinic hours to accommodate young women's schedules.

4.1.3 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 2.68 (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).

As part of the study, a Participant Tracking System (PTS; see also Section 3.2.4) 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.

4.2 Sample Characteristics

4.2.1 Study Sample

Baseline characteristics of the overall *SSI* study sample and for each replication site are presented in Exhibit 4.1. At baseline, the young women in the study sample were, on average, 17.2 years old. 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).

Across all three replication sites, more than 75 percent of youth lived with one or both biological parents. Overall, less than half said that they felt very close to and cared for by their mothers, and less than a third reported they felt close to and cared for by their fathers. Across all three sites, more than three-quarters had ever used alcohol, more than two-thirds had ever used marijuana, and just over half had ever smoked cigarettes. 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.

Given the eligibility criteria for the program, it is not surprising that most of the participants were sexually active and intended to be sexually active in the next twelve months. More than 90 percent of the sample had been sexually active (engaged in sexual intercourse, oral sex, and/or anal sex) at the time of study enrollment, 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. More than 80 percent of participants in all three replications sites intended to have sexual intercourse in the next twelve 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 twelve months.

A majority of participants reported receiving information on birth control methods and where to obtain birth control, as well as information on STIs in the year prior to the study. Fewer participants reported receiving information about abstinence or how to talk with your partner about sex and birth control during the same period.

Differences among sites. The site specific profiles of youth participating in the study differ from each other 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 just over 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 of 'Other' race. Almost half of the participants in PPGO were Black, and more than 25 percent were Hispanic.

In general, young women in Knox County appeared lower risk relative to their counterparts in Hennepin County and PPGO on several indicators, 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 sexual intercourse without birth control in the last 90 days. Larger proportions of young women in Knox County reported feeling close to or cared for by their parents than in the other two replication sites. Rates of sexual activity and intercourse without a condom in the last 90 days 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.

Exhibit 4.1: Baseline Characteristics of the 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	16.95	17.16	17.58	17.15	0.000 ***
Race/ethnicity[b]					
Hispanic	17.56	7.99	27.34	17.69	0.000 ***
Black	36.16	24.70	46.03	35.88	0.000 ***
White	26.14	60.77	21.50	32.95	0.000 ***
Other	20.14	6.54	5.14	13.49	0.000 ***
Family structure and relationships					
Lives with biological parent/s	81.19	76.90	75.82	78.94	0.040 *
Feels very close to and cared for by father	24.94	37.13	26.82	28.20	0.000 ***
Feels very close to and cared for by mother	41.71	56.27	44.71	45.76	0.000 ***
Risk behaviors					
Ever smoked cigarettes	54.46	57.46	42.29	52.25	0.000 ***
Ever drank alcohol	79.52	77.75	82.71	79.88	0.186
Ever used marijuana	73.10	59.90	62.30	67.54	0.000 ***
Knowledge, attitudes and intentions					
Knowledge of pregnancy risk	68.61	69.54	71.19	69.44	0.490
Knowledge of STI risk	67.64	67.14	70.42	68.19	0.139
Attitudes toward protection (1=least supportive, 4=most supportive)	3.25	3.30	3.22	3.25	0.014 *
Intentions to have oral sex in the next 12 months	59.96	59.95	62.91	60.65	0.552
Intentions to have sexual intercourse in the next 12 months	86.01	80.58	83.64	84.21	0.038 *
Intentions to use a condom if they were to have sexual intercourse in the next 12 months	83.92	83.74	86.45	84.48	0.435

Measure	Hennepin County	Knox County	Planned Parenthood of Greater Orlando	SSI Overall	p-value for the Test of Differences across Sitesa
Intentions to use birth control if they were to have sexual intercourse in the next 12 months	92.75	96.36	86.92	92.19	0.000 ***
Sexual behavior [c]					
Ever sexually active	94.80	90.29	94.13	93.61	0.007 **
Currently sexually active (in the last 90 days)	86.15	79.02	80.52	83.18	0.001 **
Sexual intercourse in the last 90 days	82.38	74.51	75.35	78.91	0.001 ***
Oral sex in the last 90 days	66.88	65.12	65.02	66.03	0.724
Anal sex in the last 90 days	11.68		9.39	10.97	0.209
Sexual risk [c]					
Sexual intercourse without a condom in the last 90 days	64.13	54.61	54.93	59.77	0.000 ***
Oral sex without a condom in the last 90 days	63.22	62.68	59.62	62.24	0.436
Anal sex without a condom in the last 90 days	8.97		8.22	8.74	0.648
Sexual intercourse without birth control in the last 90 days	32.01	25.97	35.21	31.39	0.013 *
Sexual intercourse with more than one partner (lifetime)	66.00	63.90	67.61	65.90	0.527
Sexual intercourse with more than five partners (lifetime)	24.18	20.24	24.59	23.37	0.231
Baseline exposure to program information[d]					
Relationships or marriage	78.57	75.00	74.30	76.74	0.139
Abstinence from sex	69.93	63.35	66.12	67.52	0.045 *
Birth control methods	90.67	87.62	75.70	86.43	0.000 ***
Where to obtain birth control	90.05	86.65	73.36	85.32	0.000 ***
Sexually transmitted infections	86.83	85.19	82.48	85.42	0.104
How to talk with partner about sex and birth control	72.23	67.96	62.15	68.86	0.001 ***
How to say no to sex	71.09	68.45	72.90	70.91	0.360
How babies are made	82.59	79.61	80.14	81.33	0.332

Source: Baseline survey administered prior to randomization.

Notes: Data in this table are based on 1,790 – 1,809 respondents (for *SSI* overall) who provided valid 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=1,780), and anal sex (n=1,385). 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] 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] 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.

[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] Questions refer to information received in the 12 months prior to the survey administration

* p< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests)

4.2.2 Comparability of the Two 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 (see Appendix Exhibit E.1).

Baseline treatment-control differences were estimated 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. At baseline, there were just two significant differences between the two groups. Young women in the treatment group reported higher baseline exposure to information on birth control methods and how to talk with a partner about sex and birth control than young women in the control group.

4.3 Program Impacts on Exposure to Sexual Health Information

In each of the replication sites, the *SSI* curriculum represented a way to provide young women with sexual health information through individualized counseling. This was intended to be over and above the sexual health information that was offered at each of the clinics as part of the routine set of services available to all young women.

Despite relatively high levels of exposure to sexual health information in the year prior to the study reported by both study groups, there is clear evidence that *SSI* significantly increased exposure to information about all eight sexual health topics queried at the short-term (9-month) follow-up (e.g., abstinence, birth control and where to obtain birth control, relationships, and how to negotiate sex), and the program effect for each of these topics was large (Exhibit 4.2).[23] Notably, the largest differences between the groups in exposure to information were for abstinence and how to say no to sex, suggesting that these messages may have been different from what was readily available outside the program (i.e., to the control group).

Exhibit 4.2: Short-term Impacts on Exposure to Sexual Health Information

Outcome	Adjusted Treatment Mean[a]	Unadjusted Control Mean	Treatment Effect[b]	p-value
Percentage of respondents that reported receiving information on the following topics:				
Relationships or marriage	86.82	71.78	15.04***	0.000
Abstinence from sex	77.92	57.75	20.17***	0.000
Birth control methods	91.43	83.52	7.90***	0.000
Where to obtain birth control	91.48	81.40	10.08***	0.000

[23] There were significant positive effects on eight out of eight measures of exposure to program information (Exhibit 4.2). In the absence of a true program impact, with eight tests and a significance criterion of $p<0.05$, the expected number of findings that would be significant by chance alone is less than one.

Outcome	Adjusted Treatment Mean[a]	Unadjusted Control Mean	Treatment Effect[b]	p-value
Sexually transmitted infections	89.98	78.96	11.02***	0.000
How to talk with partner about sex and birth control	87.74	74.06	13.68***	0.000
How to say no to sex	86.62	69.49	17.13***	0.000
How babies are made	86.81	73.08	13.72***	0.000

Source: Follow-up survey administered 9 months after baseline.

Notes: Questions refer to information received in the 12 months prior to the survey administration. Appendix D provides detailed information on measures. Results in this table are based on 1,807 – 1,809 respondents who provided valid survey responses to relevant items.

[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.

* p< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests).

4.3.1 Site-Level Differences

There was significant variation among sites in the impact of the intervention on two measures of exposure to information: 'exposure to information about birth control methods' and 'exposure to information about abstinence from sex'. In both cases, the impacts were largest in the Knox County and PPGO sites (Appendix A.1). These two programs were located in more conservative environments than Hennepin County, where perhaps there was less information about birth control methods available outside the program.

4.3.2 Subgroup Differences

Race/ethnicity. In the case of information about abstinence, although the program had a significant effect on participants from all racial/ethnic groups, the effect for White participants and those of other race/ethnicity was more than twice as large as that for Black participants and nearly twice as large as that for Hispanic participants. A similar pattern was observed for information about STIs: The impact for White participants and those of 'Other' race/ethnicity was larger than that for Hispanic participants, and no significant impact was observed for Black participants.

Age. For both older youth (age 18 and older) and younger youth (less than age 18), *SSI* increased exposure to information about relationships and marriage, how to talk with a partner about sex and birth control, how to say no to sex, and how babies are made. However, these impacts were significantly greater for older youth compared with younger youth.

4.4 Intermediate Outcomes

The *SSI* program's theory of change (see logic model in Exhibit 2.3) specifies intermediate outcomes that will influence behavior—namely, **knowledge** and understanding of sexual risk behaviors and their prevention or avoidance; **attitudes** towards sexual risk behaviors; **motivation** and **intentions** to engage in sexual activity; and refusal and negotiation **skills**.

Nine months after baseline, we find evidence that *SSI* had little impact on knowledge of pregnancy risk or STI risk, attitudes toward risky behavior, or motivation to delay childbearing. However, the program significantly and positively affected participants' attitudes toward protection, intentions, and refusal skills.

4.4.1 Knowledge

The effects of *SSI* on **knowledge of pregnancy risk** were assessed using a composite measure that combines four survey items on topics such as the effectiveness of condoms and birth control in preventing pregnancy. The effects on **knowledge of STI risk** were assessed using a composite measure that combines twelve items on STI transmission, prevention, and facts. Exhibit 4.3 shows the findings for the composite measure and the individual items.

The significant effects on exposure to information (described in Section 4.3) did not lead to corresponding significant impacts on the composite measures of knowledge of pregnancy risk and knowledge of STI risk nine months after baseline. Among the two composites and their component items, there were impacts on only one item about annual rates of STI among sexually active teens.[24] In general, participants were well informed about methods of preventing pregnancy and had general knowledge of STI facts, transmission, and prevention. In fact, study participants in both groups correctly answered 75 percent or more of the items correctly on the two composite measures of risk. This is perhaps not surprising, given that most participants were sexually active at the time of entry into the program, and might be expected to have had greater knowledge of pregnancy and STI risks.

Site/Subgroup Differences. There were no significant differences in impacts on either the composite measure of knowledge of pregnancy risk or the composite measure of knowledge of STI risk among sites (see Appendix Table A.2) or for any subgroup.

Exhibit 4.3: Short-Term Impacts on Knowledge of Pregnancy Risk and STI Risk

Outcome	Adjusted Treatment Mean[a]	Unadjusted Control Mean	Treatment Effect[b]	p-value
Knowledge of pregnancy risk (percent of items respondent answered correctly)[c]	78.53	78.26	0.27	0.817
Percent of respondents correctly answering each item:				
Used correctly, how much can birth control pills reduce pregnancy risk?	73.98	71.94	2.04	0.339
Used correctly, how much can condoms reduce pregnancy risk?	75.57	73.74	1.83	0.372
A couple that has had unprotected sex and not gotten pregnant does not have to worry about getting pregnant.	91.56	93.80	-2.24	0.083
A woman is protected from pregnancy the day she begins taking the pill.	72.99	73.57	-0.59	0.778
Knowledge of STI risk (percent of items respondent answered correctly)[d]	75.91	74.80	1.11	0.183
Percent of respondents correctly answering each item:				
Once you are infected with HIV you are infected for life	85.63	83.69	1.95	0.274
There is a vaccine to prevent girls from getting HPV	72.74	73.08	-0.35	0.871

[24] There were significant, positive effects on one out of thirteen measures of knowledge of STI risk (Exhibit 4.3). In the absence of a true program impact, with thirteen tests and a significance criterion of $p<0.05$ the expected number of findings that would be significant by chance alone is about one.

Outcome	Adjusted Treatment Mean[a]	Unadjusted Control Mean	Treatment Effect[b]	p-value
All STDs/STIs can be cured by taking medicine	74.78	77.00	-2.22	0.275
A person with an STD/STI who looks and feels healthy cannot transmit the infection to others	89.69	91.03	-1.33	0.372
Some STDs/STIs put you at greater risk of HIV	72.53	69.49	3.03	0.176
About one out of four sexually active teens gets an STD/STI every year	82.36	77.81	4.55*	0.020
You can get an STD/STI from having oral sex	91.53	89.40	2.14	0.130
Used correctly, how much can condoms decrease the risk of HIV?	56.05	55.14	0.92	0.699
You can't get infected with HIV even if you have sex only once or twice without a condom	72.33	71.78	0.55	0.801
Used correctly, how much can condoms decrease the risk of gonorrhea?	51.67	52.04	-0.37	0.875
Used correctly, how much can birth control pills decrease the risk of HIV?	80.69	78.63	2.06	0.263
Used correctly, how much can birth control pills decrease the risk of gonorrhea	81.29	78.47	2.82	0.122

Source: Follow-up survey administered 9 months after baseline.

Notes: Results in this table are based on 1,809 respondents who provided valid survey responses to relevant items.

[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] Score based on the four items below. Values shown represent the average percent of items answered correctly by respondent for each group. Alpha coefficient=0.54.

[d] Score based on the twelve items below. Values shown represent the average percent of items answered correctly by respondent for each group. Alpha coefficient=0.68.

* p< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests).

4.4.2 Attitudes

Attitudes Toward Protection. *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 protection, on average, the treatment group had slightly more positive attitudes than the control group, by three hundredths of a point on a 1-to-4 scale (an effect size of 0.09). Of the twelve items that make up the composite, there were impacts on only one item about condoms.[25]

[25] There were significant, positive effects on two out of 13 measures of attitudes toward protection (Exhibit 4.4). In the absence of a true program impact, with 13 tests and a significance criterion of p<0.05 the expected number of findings that would be significant by chance alone is about one.

Exhibit 4.4: Short-Term Impacts on Attitudes toward Protection

Outcome	Adjusted Treatment Mean[a]	Unadjusted Control Mean	Treatment Effect[b]	p-value	Effect Size[c]
Attitudes toward protection[d]	3.36	3.32	0.03*	0.050	0.09
Birth control pills should always be used if a person your age has sexual intercourse	3.42	3.37	0.04	0.245	0.06
Birth control is too much trouble to use (reverse coded)	3.29	3.29	0.01	0.806	0.01
Birth control is pretty easy to get	3.21	3.18	0.03	0.491	0.03
Birth control is important to make sex safer	3.46	3.41	0.04	0.223	0.06
Birth control has too many side effects (reverse coded)	2.66	2.64	0.02	0.563	0.03
Using birth control is morally wrong (reverse coded)	3.53	3.53	0.00	0.965	0.00
Condoms are too much trouble to use (reverse coded)	3.40	3.32	0.08*	0.037	0.10
Condoms are pretty easy to get	3.62	3.56	0.06	0.100	0.08
Condoms are important to make sex safer	3.75	3.75	0.00	0.992	0.00
Using condoms means you don't trust your partner (reverse coded)	3.58	3.56	0.01	0.703	0.02
Using condoms is morally wrong (reverse coded)	3.71	3.68	0.03	0.297	0.05
Condoms decrease sexual pleasure (reverse coded)	2.67	2.59	0.07	0.085	0.08

Source: Follow-up survey administered 9 months after baseline.

Notes: Results in this table are based on 1,800 - 1,809 respondents who provided valid survey responses to relevant items.

[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 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.

[c] 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.

[d] This construct averages responses to 12 items (shown in table) on attitudes towards condoms and birth control. Possible values for both the construct and individual items range from 1 to 4 with higher values indicating more positive attitudes toward protection (alpha coefficient=0.78).

* p< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests).

Attitudes Toward Risky Sexual Behavior. *SSI* had no statistically significant effect on the composite measure of attitudes toward risky sexual behavior nine months after baseline (Exhibit 4.5). Among the seven component items, there was an impact on only one item about having sex with someone else's partner.[26] Part of the intervention is getting young women to recognize unsafe behaviors and improve their motivation to avoid these situations. On these measures, the overwhelming majority of youth in both the treatment and control groups rejected the view that risky sexual behavior is acceptable.

[26] There were significant positive effects on one out of eight measures of attitudes toward risky sexual behavior (Exhibit 4.5). In the absence of a true program impact, with eight tests and a significance criterion of p<0.05, the expected number of findings that would be significant by chance alone is less than one.

Site/Subgroup Differences. There were no significant differences in impacts on the composite measure of attitudes toward protection among sites (see Appendix Table A.3) or for any subgroup. Nor were there any significant differences in impacts on the composite measure of attitudes toward risky sexual behavior between sites (see Appendix Table A.4) or for any subgroup.

Exhibit 4.5: Short-Term Impacts on Attitudes toward Risky Sexual Behavior

Outcome	Adjusted Treatment Mean[a]	Unadjusted Control Mean	Treatment Effect[b]	p-value
Attitudes toward risky sexual behavior (percent of items respondents agreed with)[c]	4.12	5.42	-1.30	0.061
Percent of respondents agreeing with each item:				
It's OK to have sex with someone on your first date	7.82	9.15	-1.33	0.322
It's OK to have sex with someone the same night you meet them	5.04	7.03	-1.99	0.084
It's OK to have sex with several different people in the same month	6.13	7.19	-1.06	0.382
It's okay to have sex without protection	2.08	2.45	-0.38	0.611
It's OK to have sex with someone when you know they are someone else's girlfriend/boyfriend	1.96	3.92	-1.96*	0.017
It's OK to have sex with someone if you are drunk or high	4.05	6.05	-2.00	0.059
It's OK to have sex with someone if you know they are drunk or high	1.75	2.12	-0.37	0.596

Source: Follow-up survey administered 9 months after baseline.

Notes: Results in this table are based on 1,802 respondents who provided valid survey responses to relevant items.

[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] Score based on the seven items (shown below) represents the average percent of items agreed with by respondent for each group (alpha coefficient=0.79).

* p< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests).

4.4.3 Motivation

Young women in both the treatment and control groups were highly motivated to delay childbearing. Nine months after baseline, there were no differences between the two groups on the composite measure or on any of the three individual items measuring motivation to delay childbearing. Participants in both groups indicated a belief in the importance of delaying childbearing until personal goals have been achieved (Exhibit 4.6).

Site/Subgroup Differences. There were no significant differences in impacts on the composite measure of motivation to delay childbearing among sites (see Appendix Table A.5) or for any subgroup.

Exhibit 4.6: Short-Term Impacts on Motivation to Delay Childbearing

Outcome	Adjusted Treatment Mean[a]	Unadjusted Control Mean	Treatment Effect[b]	p-value	Effect Size[c]
Motivation to delay childbearing[d]	3.76	3.73	0.03	0.309	0.05
You have goals you want to accomplish before having a child	3.71	3.70	0.02	0.583	0.03
It is important for you to finish school before you have a child	3.76	3.72	0.03	0.260	0.06
It is important to have a job and a stable income before you have a child	3.80	3.77	0.03	0.264	0.06

Source: Follow-up survey administered 9 months after baseline.

Notes: Results in this table are based on 1,802– 1,805 respondents who provided valid survey responses to relevant items.

[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 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.

[c] 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.

[d] This scale averages responses to three items (shown in table) on attitudes toward childbearing and the importance of goal setting. Possible values for both the scale and individual items range from 1 to 4 with higher values indicating greater motivation to delay childbearing (alpha coefficient=0.88).

* p< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests).

4.4.4 Intentions

Nine months after baseline, *SSI* had no overall effects on intentions to engage in sexual activity. Young women in both the treatment and control groups were equal in their expectations of engaging in sexual intercourse or oral sex in the twelve months after the survey (Exhibit 4.7). This is not surprising, given that nearly all of the participants in both groups were already sexually active (Exhibit 4.1). There were, however, large significant program effects on one of the component items, intentions to use a condom during sexual intercourse (6.6 percentage point difference) (Exhibit 4.7). [27] That is, a greater percentage of program participants reported that they intended to use a condom during sexual intercourse in the twelve months following the survey compared with participants in the control group. *SSI* had no impact on intentions to use birth control during sexual intercourse, which is unsurprising because almost all participants (over 90 percent in both groups) intended to use birth control.

Site-Level Differences. Site-level analyses showed that program impacts on intentions to use a condom during sexual intercourse varied significantly by site (Appendix Table A.6). In Hennepin County, a significantly greater percentage of program participants reported intentions to use condoms during sexual intercourse in the subsequent twelve months than control group members (a 10.5 percentage-point difference). In Knox County and PPGO, the treatment-control group differences on this outcome were

[27] There were significant positive effects on one out of four measures of intentions (Exhibit 4.7). In the absence of a true program impact, with four tests and a significance criterion of p<0.05, the expected number of findings that would be significant by chance alone is less than one.

smaller and not statistically significant. The effect on intentions to use a condom during sexual intercourse observed in the full sample was largely accounted for by Hennepin County.

Subgroup Differences. Subgroup analyses revealed a significant difference in the impacts of *SSI* on intentions to use a condom during sexual intercourse in the subsequent twelve 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 their control group counterparts (Appendix Table B.2). There were no effects on intentions to use condoms during sexual intercourse for those who were sexually inexperienced at baseline.

Exhibit 4.7: Short-Term Impacts on Intentions

Outcome	Adjusted Treatment Meana	Unadjusted Control Mean	Treatment Effectb	p-value
Intentions				
Percent of respondents reporting intentions to engage in the following behaviors in the next 12 months:				
Sexual intercourse	82.56	83.14	-0.58	0.734
Oral sex	65.95	67.05	-1.10	0.591
Use a condom if they were to have sexual intercourse	86.31	79.74	6.57***	0.000
Use birth control if they were to have sexual intercourse	92.41	91.18	1.23	0.357

Source: Follow-up survey administered 9 months after baseline.

Notes: Results in this table are based on 1,801 – 1,804 respondents who provided valid survey responses to relevant items.

[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.

* p< 0.05, ** p< 0.01, *** p< 0.001.

4.4.5 Skills

SSI had a significant positive effect on perceived refusal skills nine months after baseline. Program participants were more confident that they would be able to say no to sex under different circumstances than their control group counterparts, by one-tenth of a point on a four-point scale (an effect size of 0.17). There were no differences on perceived ability to successfully negotiate condom use with a partner (Exhibit 4.7).[28]

Site/Subgroup Differences. There were no significant differences in effects on either refusal skills or condom negotiation skills across sites (see Appendix Table A.6) or for any subgroup.

[28] There were significant positive effects on one out of two measures of skills (Exhibit 4.7). In the absence of a true program impact, with two tests and a significance criterion of p<0.05, the expected number of findings that would be significant by chance alone is less than one.

Exhibit 4.8: Short-Term Impacts on Skills

Outcome	Adjusted Treatment Mean[a]	Unadjusted Control Mean	Treatment Effect[b]	p-value	Effect Size[c]
Skills (scale score)					
Perceived refusal skills (scale score)[d]	3.45	3.34	0.10***	0.001	0.17
Perceived condom negotiation skills (scale score)[e]	3.73	3.69	0.03	0.126	0.08

Source: Follow-up survey administered 9 months after baseline.

Notes: Results in this table are based on 1,808 respondents who provided valid survey responses to relevant items.

[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 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.

[c] 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.

[d] This scale averages responses to six questions on perceived refusal skills. Possible values range from 1 to 4 with higher values indicating greater perceived skills (alpha coefficient=0.83)

[e] This scale averages responses to seven questions on perceived condom skills. Possible values range from 1 to 4 with higher values indicating greater perceived skills (alpha coefficient=0.84).

* p< 0.05, ** p< 0.01, *** p< 0.001.

4.5 Youth Sexual Behavior and Sexual Risk

After nine months, *SSI* had an overall impact on one of the two sexual risk behaviors that are the primary behavioral outcomes of interest (i.e., confirmatory outcomes, bolded in Exhibit 4.9): *currently sexually active* and *sexual intercourse without birth control*. On average, *SSI* participants were 5.8 percentage points (21 percent) less likely to report engaging in sexual intercourse without birth control than the control group. This impact is significant even after applying a multiple comparisons correction and is similar in size to the overall pooled impact for the most successful pregnancy prevention efforts (multi-component/youth development programs) found in a 2006 meta-analysis of teen pregnancy prevention efforts (Scher, Maynard, & Stagner, 2006). However, there were no program effects on current sexual activity: participants in the treatment group were no less likely to report being sexually active in the last 90 days than their control group counterparts. There were no other impacts on youth sexual behavior or sexual risk outcomes at the short-term follow-up.

4.5.1 Site-Level Differences

Site-level analyses revealed 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 treatment-control differences were observed in the Knox County or PPGO sites (Appendix A.7).

4.5.2 Subgroup Differences

Sexual experience at baseline. Nine months after baseline, significantly fewer program participants who were sexually inexperienced at baseline reported having more than one lifetime partner for sexual intercourse than their control group counterparts (a 21.4 percentage point difference; n=115) (Appendix C). This particular subgroup is interesting to look at, given the programmatic emphasis on young women

who were already sexually experienced. As stated earlier, program staff at each of the replication sites felt that characteristically, young women who were contemplating becoming sexually active were an important group to serve. For just this group, slightly less than one-third of young women in the program group became sexually active during the nine month follow-up period compared with half of the young women in the control group (Appendix C). While this treatment effect fell short of conventional levels of statistical significance (perhaps because the number of young women who were sexually inexperienced at baseline is small, providing insufficient power to detect effects), the result is suggestive of a potentially important program effect.

Age. Among older youth (age 18 and older), *SSI* significantly decreased rates of engaging in oral sex without a condom, by 7.4 percentage points on average, for treatment group members relative to control group members. There were no program effects on this outcome observed for younger youth (age less than 18). 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.

Exhibit 4.9: Short-term Impacts on Sexual Behavior and Sexual Risk

Outcome	Adjusted Treatment Mean[a]	Unadjusted Control Mean	Treatment Effect[b]	p-value
Sexual behavior (percentage responding affirmatively) [c]				
Currently sexually active (in last 90 days)	74.84	74.96	-0.11	0.954
Sexual intercourse in the last 90 days	71.29	72.18	-0.89	0.661
Oral sex in the last 90 days	59.32	60.39	-1.07	0.626
Anal sex in the last 90 days	9.13	6.13	2.99	0.051
Sexual risk (percentage responding affirmatively)				
Sexual intercourse without birth control (in last 90 days)	22.05	27.82	-5.78**	0.005
Sexual intercourse without a condom (in last 90 days)	53.66	57.45	-3.79	0.087
Oral sex without a condom (in last 90 days)	54.32	56.63	-2.31	0.299
Anal sex without a condom (in last 90 days) [c]	7.32	4.65	2.67	0.056
Sexual intercourse with more than one partner (lifetime)	70.07	71.82	-1.75	0.332
Sexual intercourse with more than five partners (lifetime)	26.35	28.86	-2.51	0.163

Source: Follow-up survey administered 9 months after baseline.

Notes: Results in this table 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). Confirmatory outcomes are bolded.

[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.

* p< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests). (For the two confirmatory outcomes statistical significance at p<0.05, p<0.01, and p<0.001 implies statistical significance at these levels after applying a Benjamini-Hochberg adjustment for multiple comparisons.)

5. Discussion

This report on short-term findings for *SSI* is the first of two impact reports. Notably, this short-term follow-up report does not analyze data on prevention of pregnancy, an important behavioral outcome that reflects the goals of OAH's Teen Pregnancy Prevention Program. A final assessment of *SSI*'s effectiveness in preventing pregnancy and reducing sexual behavior and sexual risk awaits the findings from the longer-term follow-up survey, conducted eighteen months after entry into the study. However, the short-term results presented in this report show that *SSI* did reduce sexual risk-taking behaviors among young women who were sexually active or contemplating becoming sexually active. Below, we summarize what was and was not accomplished by these ambitious efforts to replicate with fidelity a program that has been recognized as offering evidence of effectiveness.

SSI was implemented with fidelity in all three replication sites and with a broader population than the original study. With limited guidance from the program developer (who tested the program on young women hospitalized for treatment of an STI) on how to identify, recruit and retain young women from the community who were sexually active or contemplating becoming sexually active, each of the grantees independently formulated a strategy for each of these activities. In general, in each of the replication sites, nearly all participants received the initial (and arguably the most critical) session. Over the six-month period in which health educators offered booster sessions, each of the sites engaged close to two-thirds of participants in three of the four (75 percent) program sessions.

SSI had a significant positive impact on sexual risk behavior. Fewer program participants reported engaging in sexual intercourse without using birth control. This was not an impact reported in the original study. The program was also successful in addressing some potential mediators of sexual risk behavior, such as skills and intentions. Young women in the group that was offered *SSI* were more confident in their ability to say 'no' to sex and had slightly more positive attitudes toward using protection (condoms and birth control). The program also had a positive effect on intentions to use condoms when engaging in sexual intercourse. The program had no effect on knowledge, either of pregnancy risk or STI risk. Most young women in both groups were very knowledgeable about pregnancy prevention and STI facts at the outset, offering little opportunity for the program to have an effect. This is not surprising, given the program setting (health clinics) and the fact that a majority of the young women were sexually active at the time of program enrollment. The original study found increases in knowledge one month after the initial session, but no differences after that.

Site-Level Differences. The analysis found few statistically significant differences in impacts among the three replication sites. However, the evaluation plan for this study originally envisioned three equally-sized replication sites, each of which would have a large enough sample (about 1,000 study participants) for the study to detect moderately-sized impacts. In practice, the study did not achieve three equally large samples, and in only one site, Hennepin County, was the study sufficiently powered to detect moderately-sized impacts.[29] The findings for this site, which are presented in full in Appendix A, largely mirror the

[29] To mitigate the problems associated with multiple hypothesis testing, the evaluation team committed to discussing such site-level impacts if and only if significant differences were found across sites. The basic idea behind this strategy is that it is only credible to state that the program had an impact in a site – but not in another

pooled findings across all three sites, albeit with somewhat larger impact estimates and a greater number of statistically significant findings. In Hennepin County, there were large impacts on every measure of exposure to program information except for information about birth control methods. In this replication site, there were no impacts on any measures of knowledge, attitudes toward protection, motivation to delay childbearing. In Hennepin County, *SSI* did have a significant positive impact on one out of eight measures of attitudes risky sexual behavior (reducing the attitude that it is "OK to have sex with someone the same night you meet them"), and it had a positive impact on one out of six measures of intentions and skills (intention to use a condom for sexual intercourse). Notably, compared with their control group counterparts, program participants in Hennepin County were significantly less likely to engage in oral sex in the last 90 days, less likely to have oral sex without a condom, and less likely to have sexual intercourse without birth control. In addition, program participants in Hennepin County had significantly fewer lifetime partners (for sexual intercourse). However, program participants were also more likely than the control group to report having anal sex in the last 90 days.

Subgroup Differences. There were significant and important impacts by subgroup. For the majority of participants who were sexually experienced at baseline, there was a significant program impact on intentions to use condoms. For the small group of participants who were not sexually experienced at baseline (n=124), there were positive program effects on behavior. Interestingly, the program appears to have delayed sexual initiation, meaning that, after nine months, fewer *SSI* participants who were sexually inexperienced at baseline (n=84) had been sexually active than participants in the control group (n=40). In this same group of young women who were sexually inexperienced at baseline, those who participated in *SSI* were significantly less likely to report having more than one lifetime partner (for sexual intercourse) if they became sexually active in the nine month follow-up period.

The replication sites discussed here served a different population than the original study. In that study, enrollment coincided with a diagnosis of an STI, and the intervention was individualized to each participant's stage of readiness to change. Participants enrolled in the current study were not necessarily seeking treatment for an STI, but were presumably engaging in behaviors that potentially put them at risk for infection or pregnancy. Therefore, this was an opportune time to intervene. The practical implications of serving this broader population meant that determining readiness to change upon enrollment was more complicated. That is, there were clearly some young women who were engaging in sexual behaviors but at lower risk than others. However, each of the replication sites felt that these program messages were important regardless of where the young women were in terms of readiness for change. In each of the replication sites, a small percentage of young women were not sexually experienced at the time of program entry, but had come to the clinic with the intention of engaging in sexual behaviors in the immediate future. Arguably, like the young women in the original study, these inexperienced females, on the brink of being sexually active, were at a 'teachable moment.'

These early findings provide reason for optimism. The behavioral effects found in this short-term analysis are what we might reasonably expect, given that the majority of the study participants were sexually active and at relatively high risk. The subgroup findings for those not sexually experienced

---

site – if there is a significant difference in impacts between the sites. The presentation of findings in the results section of this report adheres to this plan.

at baseline, along with the consistent pattern of findings on intentions and skills overall, suggest that effects may be sustained or perhaps even amplified over time.

The findings also suggest some implications for clinical practice. Motivational interviewing, while 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 3 out of the 4 sessions). Faced with the challenge of changing established behaviors, the technique produced changes in the skills and intentions that ultimately lead to the necessary actions. Although it is hard to disentangle the roles of motivational interviewing and repeated contact, it is likely that the follow-up sessions helped to reinforce the participants' intentions to take action, and to support the change process.

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, replicated with different populations, and in different settings. The three program models were intentionally selected to maximize what could be learned about different strategies and to begin to address identified gaps in the teen pregnancy prevention research. This report, part of a larger set of reports on replications of evidence-based program models, provides important information on the early effectiveness of *SSI*.

At the very least, it suggests that strong replications of the program can have impacts on intermediate outcomes, thought to lead to behavior change, that are comparable to and even exceeding those reported by the evaluator and program developer in the earlier study.

Short-term impact reports on the other two models (*¡Cuídate!* and *Reducing the Risk*) will shed additional light on our understanding of different strategies for addressing youth risk behavior and promoting healthy choices for youth. The three final reports on longer-term outcomes will provide more comprehensive evidence on the effectiveness of these models on sexual risk-taking behaviors and their consequences.

References

American Academy of Pediatrics. (2013). Policy statement: Condom use by adolescents. doi: 10.1542/peds.2013-2831.

Bloom, H. S. (2006). *The core analytics of randomized experiments for social research. MDRC Working papers on research methodology*. Available online at: http://www.mdrc.org/publications/437/full.pdf

CDC. (2012). *Estimated HIV incidence among adults and adolescents in the US 2007-2010. HIV Surveillance Report vol. 17.* Atlanta, GA: Author.

CDC-P. (2004). *HIV surveillance report, 2003, vol.15.* Atlanta, GA: Author.

Constantine, J., Player D., Silva, T., Hallgren, K., Grider, M., and 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.

Gleason, P., Clark, M., Tuttle, C. and 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. and 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. and 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.

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

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.

Scher, L., Maynard, R., & Stagner, M. (2006). Interventions intended to reduce pregnancy-related outcomes among adolescents. Campbell Systematic Reviews; 2(12). Table 5, p. 25.

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

Schochet P.Z. (2008b). 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.

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 Education, *Institute of Education Sciences*, What Works Clearinghouse. Policies and Procedures Handbook, version 3.0 Retrieved from http://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v3_0_standards_handbook.pdf

U.S. Department of Health and Human Services. (2015). HHS Teen Pregnancy Prevention Evidence Review [website]. http://tppevidencereview.aspe.hhs.gov/

U.S. Department of Health and Human Services. (2010). Teenage pregnancy prevention: Replication of evidence-based programs. Funding opportunity announcement and application instructions. Office of Adolescent Health, Office of Public Health and Science.

U.S. Census Bureau. (2009). CB09-75. Washington, DC: Author.

U.S. Census Bureau. (2012). CB12-FF19. Washington, DC: Author.

Appendix A: Site-Level Impacts

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 either 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 A, 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 impacts are discussed in the main text, as we only interpret site-specific impacts when a significant difference across sites is found.

Exhibit A.1: Impacts on Exposure to Program Information at Short-term Follow-up, by Site

Outcome	Hennepin County (n=968)				Knox County (n=413)				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	Adj. T Mean[b]	Unadj. C Mean	T Effect[c]	p-value	Adj. T Mean[b]	Unadj. C Mean	T Effect[c]	p-value	
Percentage of respondents that reported receiving information on the following topics[d]:													
Relationships or marriage	87.00	71.73	15.27 ***	0.000	86.66	68.12	18.54 ***	0.000	86.50	75.34	11.16 **	0.004	0.404
Abstinence from sex	76.32	59.88	16.44 ***	0.000	82.58	52.90	29.68 ***	0.000	76.92	57.53	19.39 ***	0.000	0.047 *
Birth control methods	92.29	88.75	3.54	0.089	91.96	78.26	13.70 ***	0.000	88.81	76.71	12.10 ***	0.000	0.008 **
Where to obtain birth control	93.65	85.11	8.54 ***	0.000	92.21	78.99	13.22 ***	0.000	85.85	75.34	10.51 **	0.001	0.488
Sexually transmitted infections	90.58	79.64	10.94 ***	0.000	90.04	76.81	13.23 ***	0.000	88.53	79.45	9.08 **	0.007	0.686
How to talk with partner about sex and birth control	88.34	74.77	13.57 ***	0.000	87.42	73.91	13.51 ***	0.000	86.70	72.60	14.10 ***	0.000	0.991
How to say no to sex	87.65	69.91	17.74 ***	0.000	84.22	70.29	13.93 ***	0.000	86.65	67.81	18.84 ***	0.000	0.627
How babies are made	87.81	77.51	10.30 ***	0.000	86.69	68.84	17.85 ***	0.000	84.50	67.12	17.38 ***	0.000	0.130

Source: Follow-up survey administered 9 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. 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] Refers to information received in the 12 months prior to the survey administration.

p< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests).

Exhibit A.2: Impacts on Knowledge of Pregnancy Risk and STI Risk by Site

Outcome	Hennepin County (n=968)				Knox County (n=413)				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	Adj. T Mean[b]	Unadj. C Mean	T Effect[c]	p-value	Adj. T Mean[b]	Unadj. C Mean	T Effect[c]	p-value	
Knowledge of pregnancy risk (percent of items respondents answered correctly) [d]	78.91	77.81	1.10	0.485	80.61	83.33	-2.72	0.258	75.76	74.49	1.27	0.589	0.367
Percentage of respondents correctly answering each item:													
Used correctly, how much can birth control pills reduce pregnancy risk?	74.56	72.95	1.61	0.582	76.17	76.09	0.08	0.986	70.62	65.75	4.87	0.263	0.726
Used correctly, how much can condoms reduce pregnancy risk?	75.64	73.25	2.39	0.396	75.79	78.26	-2.47	0.565	75.27	70.55	4.72	0.260	0.468
A couple that has had unprotected sex and not gotten pregnant does not have to worry about getting pregnant.	90.62	93.31	-2.69	0.128	93.52	95.65	-2.13	0.429	91.82	93.15	-1.33	0.614	0.911
A woman is protected from pregnancy the day she begins taking the pill.	74.87	71.73	3.14	0.270	76.91	83.33	-6.42	0.140	65.19	68.49	-3.30	0.437	0.141
Knowledge of STI risk (percent of items respondents answered correctly) [e]	75.62	75.35	0.27	0.811	77.46	75.30	2.16	0.217	75.04	73.06	1.98	0.245	0.561
Percentage of respondents correctly answering each item:													
Once you are infected with HIV you are infected for life	83.75	82.37	1.38	0.570	88.72	83.33	5.39	0.147	86.90	86.99	-0.09	0.979	0.541
There is a vaccine to prevent girls from getting HPV	72.31	74.47	-2.16	0.461	75.63	73.91	1.72	0.701	70.89	69.18	1.71	0.695	0.664
All STD/STIs can be cured by taking medicine	72.31	75.99	-3.68	0.186	78.76	81.16	-2.40	0.571	76.56	75.34	1.22	0.768	0.616
A person with an STD/STI who looks and feels healthy cannot transmit the infection to others	89.46	91.49	-2.03	0.321	91.95	93.48	-1.53	0.623	88.08	87.67	0.41	0.894	0.800
Some STDs/STIs put you at greater risk of HIV	69.99	66.26	3.73	0.224	76.15	73.19	2.96	0.527	74.84	73.29	1.55	0.734	0.924
About 1 out of 4 sexually active teens gets an STD/STI every year	80.59	76.90	3.69	0.169	86.89	83.33	3.56	0.385	82.08	74.66	7.42	0.064	0.713

Outcome	Hennepin County (n=968)				Knox County (n=413)				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	Adj. T Mean[b]	Unadj. C Mean	T Effect[c]	p-value	Adj. T Mean[b]	Unadj. C Mean	T Effect[c]	p-value	
You can get an STD/STI from having oral sex	91.40	88.45	2.95	0.127	92.31	87.68	4.63	0.116	91.08	93.15	-2.07	0.473	0.219
Used correctly, how much can condoms decrease the risk of HIV?	59.70	57.75	1.95	0.549	48.18	53.62	-5.44	0.271	55.40	50.68	4.72	0.328	0.305
You can get HIV even if you unprotected sex only 1 or 2 times	70.71	73.86	-3.15	0.291	74.00	73.19	0.81	0.859	74.29	65.75	8.54	0.055	0.092
Used correctly, how much can condoms decrease the risk of gonorrhea?	54.25	55.02	-0.77	0.813	48.98	42.75	6.23	0.209	48.31	54.11	-5.80	0.230	0.217
Used correctly, how much can birth control pills decrease the risk of HIV?	81.27	81.46	-0.19	0.940	84.21	78.26	5.95	0.121	75.92	72.60	3.32	0.376	0.379
Used correctly, how much can birth control pills decrease the risk of gonorrhea?	82.58	80.24	2.34	0.350	84.38	79.71	4.67	0.221	75.42	73.29	2.13	0.567	0.858

Source: Follow-up survey administered 9 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. 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] Score based on the four items below. Values represent the average percent of items answered correctly by respondent for each group.

[e] Score based on the 12 items below. Values shown represent the average percent of items answered correctly by respondent for each group.

* p< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests).

Exhibit A.3: Impacts on Attitudes Toward Protection by Site

Outcome	Hennepin County (n=968)					Knox County (n=413)					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]	
Attitudes toward protection [e]	3.36	3.34	0.02	0.520	0.04	3.35	3.34	0.01	0.740	0.03	3.37	3.27	0.10**	0.006	0.25	0.125
Birth control pills should always be used if a person your age has sexual intercourse	3.44	3.38	0.06	0.266	0.07	3.47	3.48	-0.02	0.852	-0.02	3.32	3.25	0.07	0.367	0.09	0.694
Birth control is too much trouble to use (reverse)	3.28	3.29	0.00	0.944	0.00	3.35	3.38	-0.03	0.687	-0.04	3.27	3.19	0.08	0.316	0.10	0.572
Birth control is pretty easy to get	3.25	3.25	0.01	0.898	0.01	3.27	3.27	0.01	0.937	0.01	3.05	2.97	0.09	0.252	0.11	0.650
Birth control is important to make sex safer	3.46	3.46	0.00	0.924	0.01	3.47	3.47	0.00	0.987	0.00	3.42	3.25	0.17*	0.019	0.23	0.134
Birth control has too many side effects (reverse)	2.62	2.66	-0.04	0.439	-0.05	2.74	2.70	0.04	0.650	0.04	2.67	2.52	0.15	0.058	0.18	0.130
Using birth control is morally wrong (reverse)	3.53	3.55	-0.02	0.582	-0.03	3.55	3.57	-0.02	0.744	-0.03	3.53	3.47	0.06	0.291	0.10	0.465
Condoms are too much trouble to use (reverse)	3.36	3.33	0.03	0.551	0.04	3.34	3.30	0.04	0.571	0.05	3.53	3.32	0.21**	0.005	0.27	0.113
Condoms are pretty easy to get	3.64	3.59	0.05	0.303	0.07	3.59	3.51	0.08	0.227	0.12	3.59	3.54	0.04	0.515	0.06	0.892
Condoms are important to make sex safer	3.74	3.76	-0.02	0.561	-0.04	3.72	3.70	0.02	0.742	0.03	3.81	3.77	0.03	0.571	0.06	0.681
Using condoms means you don't trust your partner (reverse)	3.58	3.60	-0.02	0.653	-0.03	3.52	3.50	0.02	0.752	0.03	3.61	3.53	0.08	0.252	0.11	0.480
Using condoms is morally wrong (reverse)	3.73	3.71	0.01	0.689	0.03	3.67	3.64	0.04	0.526	0.06	3.71	3.66	0.05	0.359	0.09	0.854
Condoms decrease sexual pleasure (reverse)	2.64	2.55	0.08	0.171	0.08	2.52	2.51	0.01	0.942	0.01	2.88	2.76	0.12	0.156	0.13	0.637

Source: Follow-up survey administered 9 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. 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] This construct averages responses to 12 items (shown in table) on attitudes towards condoms and birth control. Possible values range from 1 to 4 with higher values indicating more positive attitudes toward protection

* p< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests).

Exhibit A.4: Impacts on Attitudes Toward Risky Sexual Behavior by Site

Outcome	Hennepin County (n=964)				Knox County (n=412)				Planned Parenthood of Greater Orlando (n=426)				p-value for the Test of Differences Across Sites[a]
	Adj. T Mean[b]	Unadj. C Mean	T Effect[c]	p-value	Adj. T Mean[b]	Unadj. C Mean	T Effect[c]	p-value	Adj. T Mean[b]	Unadj. C Mean	T Effect[c]	p-value	
Attitudes toward risky sexual behavior (percent of items respondents agreed with) [d]	3.92	5.64	-1.72	0.070	3.18	4.90	-1.72	0.236	5.42	5.38	0.04	0.978	0.555
Percentage of respondents agreeing with each item:													
It's OK to have sex with someone on your first date	6.73	9.12	-2.39	0.195	8.03	8.76	-0.73	0.795	10.05	9.59	0.46	0.868	0.670
It's OK to have sex with someone the same night you meet them	4.06	7.60	-3.54*	0.025	5.43	6.57	-1.14	0.634	6.82	6.16	0.66	0.778	0.305
It's OK to have sex with several different people in the same month	5.07	7.90	-2.83	0.088	5.25	5.11	0.14	0.955	9.26	7.53	1.73	0.485	0.268
It's okay to have sex without protection	2.50	3.04	-0.54	0.597	0.73	1.46	-0.73	0.635	2.37	2.05	0.32	0.830	0.863
It's OK to have sex with someone when you know they are someone else's girlfriend/boyfriend	1.91	3.34	-1.43	0.203	0.12	2.92	-2.80	0.102	3.81	6.16	-2.35	0.161	0.772
It's OK to have sex with someone if you are drunk or high	4.64	6.38	-1.74	0.228	2.38	6.57	-4.19	0.058	4.32	4.79	-0.47	0.829	0.466
It's OK to have sex with someone if you know they are drunk or high	2.58	2.13	0.45	0.641	0.35	2.92	-2.57	0.079	1.28	1.37	-0.09	0.952	0.219

Source: Follow-up survey administered 9 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. 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] Score based on the seven items (shown below) represents the average percent of items agreed with by respondent for each group.

* p< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests).

Exhibit A.5: Impacts on Motivation to Delay Childbearing by Site

Outcome	Hennepin County (n=966)					Knox County (n=412)					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]	
Motivation to delay childbearing[e]	3.75	3.73	0.03	0.476	0.05	3.76	3.73	0.03	0.620	3.76	3.76	3.74	0.03	0.594	0.05	0.999
You have goals you want to accomplish before having a child	3.71	3.68	0.03	0.518	0.04	3.71	3.70	0.01	0.846	3.71	3.72	3.73	0.00	0.976	0.00	0.925
It is important for you to finish school before you have a child	3.75	3.74	0.01	0.845	0.01	3.75	3.67	0.08	0.209	3.75	3.78	3.73	0.05	0.431	0.08	0.617
It is important to have a job and a stable income before you have a child	3.80	3.76	0.04	0.275	0.07	3.81	3.81	0.00	0.939	3.81	3.79	3.75	0.04	0.463	0.07	0.785

Source: Follow-up survey administered 9 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. 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] This scale averages responses to 3 items (shown in table) on attitudes toward childbearing and the importance of goal setting. Possible values range from 1 to 4 with higher values indicating greater motivation to delay childbearing.

* p< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests).

Exhibit A.6: Impacts on Intentions and Skills 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]	
Intentions																
Percentage of respondents reporting intentions to engage in the following behaviors in the next 12 months:																
Sexual intercourse	84.57	85.98	-1.41	0.549		78.59	81.75	-3.16	0.380		81.79	78.08	3.71	0.289		0.342
Oral sex	65.64	68.39	-2.75	0.326		66.42	66.67	-0.25	0.954		66.15	64.38	1.77	0.671		0.650
Use a condom if they were to have sexual intercourse	86.82	76.29	10.53 ***	0.000		81.23	83.21	-1.98	0.592		90.14	84.25	5.89	0.102		0.018
Use birth control if they were to have sexual intercourse	94.00	91.19	2.81	0.126		92.87	90.51	2.36	0.398		88.45	91.78	-3.33	0.222		0.156
Skills																
Perceived refusal skills (scale score) [e]	3.43	3.36	0.08	0.064	0.13	3.45	3.36	0.10	0.123	0.16	3.46	3.30	0.17 **	0.007	0.27	0.470
Perceived condom negotiation skills (scale score) [f]	3.73	3.70	0.03	0.318	0.07	3.74	3.69	0.05	0.239	0.12	3.72	3.70	0.02	0.623	0.05	0.870

Source: Follow-up survey administered 9 months after baseline.

[a] This column shows the statistical result for the test of differences in the treatment effect across 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] This scale averages responses to 6 questions on perceived refusal skills. Possible values range from 1 to 4 with higher values indicating greater perceived skills.

[f] This scale averages responses to 7 questions on perceived condom skills. Possible values range from 1 to 4 with higher values indicating greater perceived skills.

* p< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests).

Exhibit A.7: Impacts on Sexual Behavior and Sexual Risk 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. T Mean[b]	Unadj. C Mean	T Effect[c]	p-value	Adj. T Mean[b]	Unadj. C Mean	T Effect[c]	p-value	Adj. T Mean[b]	Unadj. C Mean	T Effect[c]	p-value	
Sexual behavior (percentage responding affirmatively) [d]													
Currently sexually active (in last 90 days)	76.19	78.66	-2.47	0.361	74.03	68.61	5.42	0.188	72.41	72.60	-0.19	0.963	0.277
Sexual intercourse in the last 90 days	72.92	76.22	-3.30	0.236	70.74	66.42	4.32	0.308	67.93	68.49	-0.56	0.893	0.322
Oral sex in the last 90 days	57.97	64.02	-6.05 *	0.044	61.22	55.47	5.75	0.209	60.30	56.85	3.45	0.439	0.050*
Anal sex in the last 90 days	9.59	5.50	4.09 *	0.027					8.10	7.53	0.57	0.835	0.287
Sexual risk (percentage responding affirmatively)													
Sexual intercourse without birth control (in last 90 days)	21.43	29.27	-7.84 **	0.005	17.58	23.36	-5.78	0.173	27.57	28.77	-1.20	0.772	0.412
Sexual intercourse without a condom (in last 90 days)	57.43	62.80	-5.37	0.077	49.94	53.28	-3.34	0.470	48.60	49.32	-0.72	0.873	0.689
Oral sex without a condom (in last 90 days)	53.20	60.06	-6.86 *	0.024	57.63	54.74	2.89	0.533	53.49	50.68	2.81	0.535	0.092
Anal sex without a condom (in last 90 days)	7.75	4.59	3.16	0.060					6.38	4.79	1.59	0.524	0.601
Sexual intercourse with more than one partner (lifetime)	25.85	30.09	-4.24	0.085	29.55	29.69	-0.14	0.971	24.49	25.35	-0.86	0.814	0.581
Sexual intercourse with more than five partners (lifetime)	21.43	29.27	-7.84 **	0.005	17.58	23.36	-5.78	0.173	27.57	28.77	-1.20	0.772	0.412

Source: Follow-up survey administered 9 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 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< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests).

Appendix B: Subgroup Impacts

Exhibit B.1: Impacts on Exposure to Program Information at Short-term Follow-up, by Subgroup

	Treatment Effect [a]	p-value [b]
Received information about relationships or marriage (percentage of respondents)		
Subgroup: Respondent Age		
Respondent less than age 18 (n=966)	11.03***	0.000
Respondent age 18 or older (n=843)	19.28***	0.000
Received information about abstinence (percentage of respondents)		
Subgroup: Respondent Race		
Hispanic (n=320)	16.17**	0.001
Black (n=649)	10.85**	0.003
White (n=596)	30.05***	0.000
Other (n=244)	26.61***	0.000
Received information about STIs (percentage of respondents)		
Subgroup: Respondent Race		
Hispanic (n=320)	11.50**	0.003
Black (n=649)	4.84	0.080
White (n=596)	16.20***	0.000
Other (n=244)	14.16**	0.002
Received information about how to talk with partner about sex and birth control (percentage of respondents)		
Subgroup: Respondent Age		
Respondent less than age 18 (n=966)	6.95**	0.005
Respondent age 18 or older (n=843)	21.13***	0.000
Received information about how to say no to sex (percentage of respondents)		
Subgroup: Respondent Age		
Respondent less than age 18 (n=965)	11.54***	0.000
Respondent age 18 or older (n=843)	23.30***	0.000
Received information about how babies are made (percentage of respondents)		
Subgroup: Respondent Age		
Respondent less than age 18 (n=965)	7.40**	0.003
Respondent age 18 or older (n=842)	20.82***	0.000

Source: Follow-up survey administered 9 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<0.05). For example, a test result indicated that the treatment effect on learning about abstinence was significantly different across racial/ethnic groups.

[a] This column shows the estimated treatment effect (treatment/control difference in the percent reporting receiving information) 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< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests).

Exhibit B.2: Impacts on Intentions, by Subgroup

	Treatment Effect [a]	p-value [b]
Intentions to use a condom if they were to have sexual intercourse (percentage of respondents)		
Subgroup: Respondent sexual experience at baseline		
Never sexually active at baseline (n=125)	-6.77	0.334
Ever sexually active at baseline (n=1,679)	7.50***	0.000

Source: Follow-up survey administered 9 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<0.05). For example, a test result indicated that the treatment effect on intention to use a condom or birth control if they were to have sexual intercourse was significantly different between the sexually experienced at baseline subgroups.

[a] This column shows the estimated treatment effect (treatment/control difference in the percent reporting intentions) 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< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests).

Exhibit B.3: Impacts on Sexual Behavior and Sexual Risk, 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	0.294
Respondent age 18 or older (n=839)	-5.52	0.060
Oral sex without a condom in the last 90 days		
Subgroup: Respondent Age		
Respondent less than age 18 (n=963)	2.05	0.507
Respondent age 18 or older (n=838)	-7.43*	0.021
Sexual intercourse with more than one lifetime sexual partner		
Subgroup: Respondent sexual experience at baseline		
Never sexually active at baseline (n=115)	-21.42*	0.028
Ever sexually active at baseline (n=1,620)	-0.50	0.790
Subgroup: Respondent Race		
Hispanic (n=309)	-9.75*	0.020
Black (n=612)	1.03	0.736
White (n=580)	2.53	0.422
Other (n=234)	-8.31	0.097

Source: Follow-up survey administered 9 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 ($p<0.05$). 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.

* $p< 0.05$, ** $p< 0.01$, *** $p< 0.001$ (two-tailed tests).

Appendix C: Behavioral Impacts for Sexually Inexperienced Youth

Exhibit C.1: Short-term Impacts 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 behavior (percentage responding affirmatively) [c]				
Sexual initiation	29.10	50.00	-20.90	0.097
Currently sexually active (in last 90 days)	22.06	32.50	-10.44	0.338
Sexual intercourse in the last 90 days	18.07	27.50	-9.43	0.329
Oral sex in the last 90 days	11.69	22.50	-10.81	0.235
Anal sex in the last 90 days	3.18	7.14	-3.97	0.591
Sexual risk (percentage responding affirmatively)				
Sexual intercourse without birth control (in last 90 days)	5.51	10.00	-4.49	0.506
Sexual intercourse without a condom (in last 90 days)	13.00	12.50	0.50	0.950
Oral sex without a condom (in last 90 days)	12.79	20.00	-7.21	0.402
Anal sex without a condom (in last 90 days)	2.38	3.57	-1.19	0.777
Sexual intercourse with more than one partner (lifetime)	9.14	30.56	-21.42*	0.028
Sexual intercourse with more than five partners (lifetime)	6.38	5.56	0.82	0.869

Source: Follow-up survey administered 9 months after baseline.

Notes: 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< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests).

Appendix D: Measures

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

- Exposure to information;
- Intermediate outcomes; and
- Youth sexual behavior.

Measures in the first category (*exposure to program information*) reflect receipt of sexuality education and reproductive health information. These provide insight into *SSI's* success in reaching youth. Measures of *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. Measures of *youth sexual behavior* include measures of sexual activity and sexual risk behavior (e.g., unprotected sexual activity). In the sections that follow, we describe each category by defining constituent measures and their construction.

D.1 Exposure to Program Information

To assess whether *SSI* increased exposure to information on sexual health, contraception, and STI transmission and prevention, at the short-term follow-up, we asked youth about their receipt of sexuality education and reproductive health information. [30] On the survey, they responded to a series of questions asking about their exposure to information about: (a) relationships or marriage; (b) abstinence from sex; (c) birth control methods; (d) where to obtain birth control; (e) STIs; (f) how to talk with a partner about sex and birth control; (g) how to say no to sex; and (h) how babies are made. For each, youth were asked whether they had "received information or learned about" the topic in the 12 months prior to survey administration. Responses were coded in a binary fashion, as 1="yes" and 0="no."

D.2 Intermediate Outcomes

Intermediate outcomes are those expected to portend changes in behavior. At the short-term follow-up, 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 (where 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.

[30] At baseline, before random assignment, youth were asked these same questions about the twelve month period preceding the study. The reference period for the follow-up survey included the period in which treatment group members were offered the intervention (and controls were not).

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 Exhibit D.2.1.

- ***Knowledge of pregnancy risk*** is a composite measure that is the mean (multiplied by 100) of four binary variables regarding the extent to which contraceptive methods can prevent pregnancy and circumstances under which pregnancy is possible (See Exhibit D.2.1 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 of twelve binary variables (multiplied by 100) pertaining to STI prevention, transmission, and treatment (See Exhibit D.2.1 for coding and other details.) Scores on this scale range from 0 to 100 and represent the percentage of correct answers across the twelve items. Higher values indicate more accurate knowledge.

Exhibit D.2.1: 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?	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.
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.	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.
Knowledge of STI Risk (12 items)	
You can't get infected with HIV if you have sex only once or twice w/o 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.	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 true and 0 indicates they were sure or thought the statement was false or did not know.
There is a vaccine to prevent girls from getting HPV.	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 true and 0 indicates they were sure or thought the statement was false or did not know.
All STDs/STIs can be cured by taking medicine.	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 person with an STD/STI who looks and feels healthy cannot transmit the infection to others.	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.
Some STDs/STIs out you at greater risk of HIV.	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 true and 0 indicates they were sure or thought the statement was false or did not know.
About 1 out of 4 sexually active teens gets an STD/STI every year.	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 true and 0 indicates they were sure or thought the statement was false or did not know.

Component Items	Coding
You can get an STD/STI from having oral sex.	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 true and 0 indicates they were sure or thought the statement was false or did not know.
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 ("a lot") was coded as 1 and all other responses were coded as 0.
Used correctly, how much can condoms decrease the risk of gonorrhea?	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 birth control pills 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 birth control pills decrease the risk of gonorrhea?	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.

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 D.2.2.

- ***Attitudes toward protection*** is a composite measure that is the mean of responses to twelve items about the importance of using condoms and/or birth control during sexual activity. (See Exhibit D.2.2 for coding and other details.) Scores on this scale represent 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 demonstrates acceptable internal consistency reliability (α=0.78). [31]
- ***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 risky sexual behaviors. (See Exhibit D.2.2 for coding and other details.) Scores on this scale range from 0 to 100 and represent the percent of items agreed with: Higher values reflect more support for risky behavior. The measure demonstrates acceptable internal consistency reliability (α=0. 79).

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

Exhibit D.2.2: 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." We reverse coded this item so that higher values indicate more positive attitudes toward birth control.
Birth control is too much trouble to use.	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 birth control.
Birth control is pretty easy to get.	Youth expressed their agreement with this statement, responding on a scale from 1= "Strongly agree" to 4="Strongly disagree." We reverse coded this item so that higher values indicate more positive attitudes toward birth control.
Birth control is important to make sex safer.	Youth expressed their agreement with this statement, responding on a scale from 1= "Strongly agree" to 4="Strongly disagree." We reverse coded this item so that higher values indicate more positive attitudes toward birth control.
Birth control has too many side effects.	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 birth control.
Using birth control is morally wrong.	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 birth control.
Condoms are too much trouble to use.	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.
Condoms are pretty easy to get.	Youth expressed their agreement with this statement, responding on a scale from 1= "Strongly agree" to 4="Strongly disagree." We reverse coded this item so that higher values indicate more positive attitudes toward condoms.
Condoms are important to make sex safer.	Youth expressed their agreement with this statement, responding on a scale from 1= "Strongly agree" to 4="Strongly disagree." We reverse coded this item so that higher values indicate more positive attitudes toward condoms.
Using condoms means you don't trust your partner.	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.
Using condoms is morally wrong.	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.
Condoms decrease sexual pleasure.	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.
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.	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.

Component Items	Coding
It's OK to have sex with several different people in the same month.	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 without protection.	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 when you know they are someone else's girlfriend/boyfriend.	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 if you are drunk or high.	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 if you know they are drunk or high.	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.

Motivation

The short-term survey 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 Exhibit D.2.3 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 (α=0.88).

Exhibit D.2.3: 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.	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 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.	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.

Intentions

We used the four items presented in Exhibit D.2.4 to examine impacts on youth's intended or anticipated sexual behavior in the coming year.

Exhibit D.2.4: 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?	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.
If you have sexual intercourse in the next year, do you intend to use birth control?	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.
If you have sexual intercourse in the next year, do you intend to use a condom?	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.

Skills

The short-term follow-up survey 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 D.2.5.

- ***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 D.2.5 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 demonstrates good internal consistency reliability (α=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 D.2.5 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 use condoms. The measure demonstrates good internal consistency reliability (α=0.84).

Exhibit D.2.5: 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?	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 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?	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 neither you nor your partner had any form of birth control?	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 have dated for a long time?	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 after you have been drinking alcohol?	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.
Condom Negotiation Skills (7 items)	
If you were going to have sex could you get or buy a condom?	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 talk about using condoms with your partner before having sex?	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 insist on using a condom if your partner didn't want to use one?	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 ask your partner to use condoms even if the two of you had sex before w/o 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?	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 ask a new partner to use 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 get a partner to use condoms, even if you're drunk or high?	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.

D.3 Youth Sexual Behavior and Sexual Risk

To understand program effects on youths' sexual behavior and sexual risk, we examined their responses to questions about their history of sexual activity, their recent sexual behavior, their recent sexual risk behavior, and the number of lifetime sexual partners. We used the eleven items presented in Exhibit D.3.1 to examine impacts on sexual behavior and sexual risk.

Exhibit D.3.1: Youth Sexual Behavior and Sexual Risk Measures

Measure	Item	Coding
Sexual Behavior		
Initiation of sexual activity	Have you ever had any of the following: sexual intercourse, oral sex or anal sex?	Youth who were not sexually active at baseline 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/or anal sex. Youth were not asked about anal sex in Knox County.
Currently sexually active (in last 90 days)	Coded from the three items that follow:	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.
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.

Measure	Item	Coding
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.
Sexual Risk		
Sexual intercourse without a condom (in last 90 days)	In the past 3 months, have you had sexual intercourse without you or 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 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.
Anal sex without a condom (in 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.
Sexual intercourse without birth control (in 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 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.

Appendix E: Supporting Tables

Exhibit E.1: Characteristics of the Analytic Sample at Baseline

Measure	Treatment Mean[a]	Control Mean	Group Difference[b]	p-value
Demographic characteristics				
Age				
Mean	17.12	17.14	-0.01	0.794
Race/ethnicity[c]				
Hispanic	16.80	19.90	-3.10	0.086
Black	35.77	35.24	0.53	0.807
White	33.58	31.65	1.93	0.335
Other	13.85	13.21	0.64	0.697
Family structure and relationships				
Lives with biological parents	78.97	78.02	0.96	0.630
Feels very close to and cared for by father	29.58	26.07	3.51	0.139
Feels very close to and cared for by mother	44.25	48.33	-4.09	0.101
Risk behaviors				
Ever smoked cigarettes	51.49	53.28	-1.79	0.465
Ever drank alcohol	78.54	82.10	-3.56	0.071
Ever used marijuana	67.29	68.03	-0.74	0.750
Knowledge[d]				
Knowledge of pregnancy risk	68.50	70.96	-2.47	0.173
Knowledge of STI risk	68.75	67.24	1.51	0.246
Attitudes[e]				
Attitudes toward protection	3.26	3.26	0.00	0.971
Intentions				
Intentions to have oral sex in the next 12 months	60.34	61.82	-1.47	0.528
Intentions to have sexual intercourse in the next 12 months	84.22	84.40	-0.18	0.917
Intentions to use a condom if they were to have sexual intercourse in the next 12 months	84.62	83.91	0.72	0.693
Intentions to use birth control if they were to have sexual intercourse in the next 12 months	92.76	91.48	1.29	0.326
Sexual Behavior[f]				
Currently sexually active (in last 90 days)	82.97	83.53	-0.56	0.761
Sexual intercourse in the last 90 days	78.70	79.08	-0.38	0.852
Oral sex in the last 90 days	66.00	66.28	-0.28	0.905
Anal sex in the last 90 days	11.52	10.68	0.84	0.637

Measure	Treatment Mean[a]	Control Mean	Group Difference[b]	p-value
Sexual Risk[f]				
Sexual intercourse without a condom in the last 90 days	59.86	59.14	0.72	0.767
Oral sex without a condom in the last 90 days	62.36	62.15	0.22	0.927
Anal sex without a condom in the last 90 days	9.58	7.91	1.68	0.297
Sexual intercourse without birth control in the last 90 days	31.31	31.47	-0.15	0.947
Sexual intercourse with more than one partner (lifetime)	66.21	67.35	-1.15	0.626
Sexual intercourse with more than 5 partners (lifetime)	23.60	23.37	0.24	0.911
Baseline exposure to program information[g]				
Relationships or marriage	77.09	76.14	0.95	0.652
Abstinence from sex	68.55	65.14	3.41	0.139
Birth control methods	87.59	83.99	3.60*	0.030
Where to obtain birth control	86.12	83.50	2.63	0.128
Sexually transmitted infections	86.16	83.66	2.50	0.154
How to talk with partner about sex and birth control	71.06	64.54	6.52**	0.004
How to say no to sex	71.50	69.93	1.56	0.488
How babies are made	82.11	80.72	1.39	0.469

Source: Baseline survey administered prior to randomization.

Notes: Results in this table are based on the analytic sample of 1,786 - 1,809 respondents who provided valid 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= 1710), and anal sex (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). Due to rounding, reported group differences may differ from differences reported between reported means for the treatment and control groups.

[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.

[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] Knowledge variables are composite scale scores representing the proportion of items answered correctly.

[e] Attitude variable is a composite scale score with higher scores indicating more positive attitudes.

[f] 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.

[g] Questions refer to information received in the 12 months prior to the survey administration.

* p< 0.05, ** p< 0.01, *** p< 0.001 (two-tailed tests).