Findings from an Innovative Teen Pregnancy Prevention Program

Evaluation of the *Crossroads* Program in Arlington, TX

Final Impact Report for

Arlington Independent School District

12/30/2015

Prepared by

Holli Slater, PhD
Diane Mitschke, PhD

University of Texas at Arlington
211 S. Cooper St, Box 19129
Arlington, TX 76019

Recommended citation:

Slater, H.M., and Mitschke, D.B. (2015). "Evaluation of the *Crossroads* Program in Arlington, TX: Findings from an Innovative Teen Pregnancy Prevention Program" Arlington, TX: University of Texas at Arlington.

Acknowledgements:

We would like to acknowledge the many people who made this study and report possible. First, the administration and staff at Arlington Independent School District who were supportive of this project from the beginning and provided the resources needed to bring this project to fruition. We would like to acknowledge the Drop Out Prevention staff that assisted with recruitment efforts and provided support to ensure youth made it to the intervention each day, without you we would not have had youth to participate in the program and study. We would like to acknowledge the *Crossroads* facilitators who ensured a quality program was implemented each and every intervention. Lastly, we want to express our appreciation for each individual youth who participated in the program and study in the midst of many other challenges. It is because of you that this program was even possible and your contributions will further inform programs like these.

Thank you to the following members of the evaluation team who contributed to this final report: Sharon Homan, who provided guidance and support for the analysis plan, Marcus Crawford, who provided support for the impact analysis, Daniel Reese and Amanda Crowe, who provided support for qualitative data collection, and lastly the numerous graduate research assistants who tirelessly searched for participants and were available at a moment's notice to administer surveys. Without you we would not have had the data to drive this project.

The Authors

This publication was prepared under Grant Number TP2AH000011 from the Office of Adolescent Health, U.S. Department of Health & Human Services (HHS). The views expressed in this report are those of the authors and do not necessarily represent the policies of HHS or the Office of Adolescent Health.

EVALUATION OF THE CROSSROADS PROGRAM IN ARLINGTON, TX: FINDINGS FROM AN INNOVATIVE TEEN PREGNANCY PREVENTION PROGRAM

I. Introduction

The U.S. Department of Health and Human Services (HHS) has placed an emphasis on reducing teen pregnancy rates. In 2010, the Teen Pregnancy Prevention (TPP) initiative was implemented to expand the breadth of evidence-based practices by funding 19 Tier 2 rigorous evaluation studies of new and innovative programs. Arlington Independent School District was funded to do this through the implementation and evaluation of an adapted version of the evidence based program *Be Proud! Be Responsible!*.

Arlington Independent School District is the 11th largest school district in the state of Texas serving more than 63,000 students a year. The need for pregnancy prevention services in this community at the time of funding was high with a trend of increasing teen pregnancies from 2008 to 2010 (Arlington Independent School District, 2011). Sexually transmitted infection (STI) rates continued to rise from 2001 to 2008 across all races and age groups within the local community (Tarrant County Health Department, 2009). The local school district, recognizing the connection between academic achievement and issues related to sexual health supported an initiative to address sexual health issues with youth in this community.

A. Introduction and study overview

Previous studies have demonstrated effectiveness of the *Be Proud! Be Responsible!* curriculum in different settings with youth between 12 and 18 years of age (Jemmott, Jemmott, Fong, & Morales, 2010; Jemmott, Jemmott, & Fong, 1998). These studies reported a reduction in unprotected sexual intercourse, increased condom use, and reduced rates of STIs (Jemmott, Jemmott, & Fong, 1998; Jemmott, Jemmott, Braverman, & Fong, 2005). This study aimed to determine if these program effects could be replicated when modified to incorporate components

aimed at developing interpersonal skills, building connections with community resources, and educational or vocational goal setting with older adolescents who are academically at risk.

Research has demonstrated a connection between academic achievement and risky sexual behavior (Frisco, 2008). Pregnancy, a consequence of risky sexual behavior, has been closely associated with lower graduation rates among young women (Perper, Peterson, & Manlove, 2010; National Campaign to Prevent Teen and Unplanned Pregnancy, 2012). In contrast, school enrollment has been seen to be a protective factor against sexually transmitted infections (Crosby, Diclemente, Wingood, Salazar, Rose, & Sales, 2007), further suggesting a relationship between academic achievement and risky sexual behavior. In 2010, Arlington Independent School District reported a four-year drop out rate of 9.2%, which was higher than the overall rate of 7% for the state of Texas at that time (Arlington Independent School District, 2008; Annie E. Casey Foundation, 2008). A strong push from the district resulted in 735 students enrolling with Drop Out Recovery Extended Services; however, 32% of these youth did not remain in the program and complete high school requirements (Arlington Independent School District, 2008; Arlington Independent School District, 2009). *Crossroads* was intended to address both the academic and sexual health education needs of older adolescents in an effort to improve outcomes.

The evaluation consists of two components: (1) an outcome study aimed at assessing the impact of the program on risky sexual behavior among older adolescents and (2) an implementation study aimed at assessing the adherence, quality, and context impacting implementation of the program.

B. Primary research question

The primary research question for this evaluation is: What is the impact of the offer to participate in *Crossroads* (intervention) relative to the graduation coach program (business as

usual counterfactual) on participants not using a condom during vaginal intercourse 3 months after the end of treatment?

C. Secondary research questions

This evaluation features two types of secondary research questions. The first set addresses the impacts explored in the primary research question at different time points. These include the following:

- What is the impact of the offer to participate in *Crossroads* (intervention) relative to the graduation coach program (business as usual counterfactual) on participants not using a condom during vaginal intercourse 6 and 12 months after the end of treatment?

An additional set of secondary research questions address impacts on other behavioral outcomes not included in the primary research question. These include the following:

- What is the impact of the offer to participate in *Crossroads* (intervention) relative to the graduation coach program (business as usual counterfactual) on participants not using a condom during anal intercourse 3, 6 and 12 months after the end of treatment?
- What is the impact of the offer to participate in *Crossroads* (intervention) relative to the graduation coach program (business as usual counterfactual) on participants not using a condom during oral intercourse 3, 6, and 12 months after the end of treatment?
- What is the impact of the offer to participate in *Crossroads* (intervention) relative to the graduation coach program (business as usual counterfactual) on participants' reported pregnancies 12 months after the end of treatment?

II. Program and comparison programming

A. Description of program as intended

Crossroads was designed to provide an open and comfortable space for youth to discuss sensitive sexual health topics. The program was structured such that a primary facilitator was

responsible for the teaching portion of curriculum in a large group setting. At various points during the intervention, participants would break off into randomly pre-assigned small groups of 10 or fewer youth. A trained facilitator engaged youth in small-group skill-building activities and guided in-depth discussions of large-group topics. In order to establish trust and facilitate engagement in the activities, groups and assigned facilitators remained consistent throughout the duration of the three days. Each facilitator was required to have a minimum of three years of experience in teaching, counseling, or social work, experience working with at risk students and experience working with teen pregnancy prevention. All facilitators were trained on how to implement the curriculum directly by the curriculum developers and received additional staff development trainings aimed at practitioners working with at risk youth.

The program dosage was intended to be 16 activities over the course of 7 hours each day for a total of 21 hours across the three-day program, but was later adapted to fit within a shorter timeframe.[1] Each day had a separate focus for program participants including building personal skills, prevention pregnancy and STIs, and identifying resources available within the community. Day 1 took place at Camp Thurman, an outdoor experiential learning adventure camp. Staff members employed by the adventure camp were the primary facilitators of activities for this day with program facilitators assisting by drawing connections between the activities and participants' sexual health. Youth participated in an outdoor challenge course and worked to identify attitudes and personal beliefs about relationships. The focus of this day was "Who am I?" and encouraged participants to become more self aware of their personal beliefs and how

[1] The program was reduced from 21 hours to 18.75 hours to accommodate the needs of the youth participating. Cohorts 1 and 2 received the longer programming, while all subsequent cohorts received the shorter programming. Components of certain activities were modified to accommodate the shorter schedule, but the number of activities offered did not change.

those beliefs impacted their decision-making. The focus of Day 2 was "Where am I Going?" and encouraged participants to see educational possibilities for themselves and understand how an unwanted pregnancy could impact those goals. This day took place at a local community college where participants participated in a college tour led by college personnel and/or students, filled out a career assessment, and discussed their potential to continue their education past high school. Participants also began to receive a modified version of the *Be Proud! Be Responsible!* curriculum facilitated by program staff, which focused on knowledge building about STIs and HIV/AIDS. Day 3 began at Mission Arlington, a community organization that provides a variety of services and resources for the local community. Participants assisted with a community service project (e.g., sorting food or clothing donations) and learned about the resources this agency could provide for them and their families, as well as how they can give back to their own community. Participants were taken to an alternate location to meet with healthcare workers from a local community clinic and learned about the sexual health resources available at the clinic. The day concluded with the completion of the *Be Proud! Be Responsible!* curriculum, facilitated by program staff, which focused on continuing knowledge building about STIs, HIV/AIDS, and pregnancy prevention. The focus of the final day was "Where Do I Get Help From Here?" and encouraged participants to identify resources in the community that can assist with meeting their goals.

The program was implemented 6 or 7 times during the academic year with each intervention occurring across three consecutive days. Facilitators continued to make contact with participants for one year following the completion of the intervention to answer questions about content and see if additional support services were needed. Data on dosage and content of these visits and/or conversations were not formally collected.

B. Description of counterfactual condition

The counterfactual experience was intended to be business as usual in which the comparison group received no portion of the treatment nor did they receive any alternative program. Arlington Independent School District is an "Abstinence Only" school district; therefore, any additional programming regarding sexual health was offered in this context to all youth. The district does offer a comprehensive program for youth who are pregnant and/or parenting, which covers some content related to sexual health. Clubs and school organizations may have offered guest speakers or mentor programming that provided some information regarding sexual health; however, none of these options provided comprehensive pregnancy and STI prevention. Some youth may have received reproductive health services from local clinics or healthcare professionals within the community. Youth assigned to either condition may have accessed any of the aforementioned services.

III. Study design

The impact study was a cohort-based randomized controlled trial design used to estimate the impact of *Crossroads* on reducing risky sexual behaviors among older adolescent youth classified as high risk for dropping out of high school. Random assignment of youth to treatment and comparison groups controls for spurious causality and allows for interpretation of outcomes attributable to the intervention rather than some other factor. The implementation study utilized mixed methods to assess the quality and context of program implementation. The following section details components of these two studies including sample recruitment, study design, data collection, outcomes, and characteristics of the analytic sample including baseline equivalence.

A. Sample recruitment

The study took place in Arlington Independent School District and included one alternative and six traditional high school campuses. Each campus offered Drop Out Prevention services and

had one graduation coach assigned to assist with academic support for both treatment and comparison group participants at their home campus and to serve as the primary recruiters for the program. Sample enrollment started in the 2011–2012 school year and ended in the 2014–2015 school year.

In order to be included in the study sample a youth must have met the following inclusion criteria:

1. Currently be enrolled in Arlington Independent School District
2. Participating in Drop Out Prevention services (i.e., working with a graduation coach for academic support)
3. 17–19 years old
4. Have previously dropped out of school and/or be considered at high risk for dropping out[2]
5. Able to read and understand English
6. Provided consent/assent
7. Able to attend a specific intervention session (3-day period)[3]

Graduation coaches identified students who met these criteria by reviewing reports generated from school district records. Any youth who met the inclusion criterion listed above

[2] To be considered at risk, a student must meet one or more the following criteria during the current academic school year: (1) not currently on grade level, (2) failed STAARS or TAKS (standardized tests used to assess students' attainment of reading, writing, math, science and social studies skills), (3) participated in an alternative education program, (4) expelled, (5) on probation, (6) homeless, (7) involved in the juvenile justice system, (8) involved in Child Protective Services, (9) limited English proficiency, (10) parenting (or expecting).

[3] Youth were required to verify their availability to participate in the upcoming intervention in order to be eligible to be randomized. This was done to maximize the likelihood of youth attendance at the intervention. Youth who indicated they were "unavailable" remained eligible for randomization in future interventions as long as they continued to meet the other inclusion criteria.

was eligible to be in the sample.[4] As the study progressed, teachers, administrators, and former program participants referred a large number of youth directly to the program. Youth were recruited and randomized on a rolling basis, with 21 cohorts of youth recruited over the three and a half years of the study.

B. Random Assignment

The research design was a randomized controlled trial with youth individually randomized to treatment or comparison. Graduation coaches recruited youth throughout the school year and reviewed the consent form[5] with participants (individually or in a group setting) and emphasized the voluntary nature of the program. Participants were assured that they would continue to receive Drop Out Prevention services regardless of their decision to consent to participate. There were no incentives for agreeing to participate and/or returning consent paperwork. All potential participants were told that consenting to participate did not guarantee they would be selected to be in either the treatment or comparison group, but gave them to opportunity to be selected at random.

Two weeks before a session, a random sample of 60 youth was drawn and graduation coaches confirmed their interest and availability to attend the session. Randomization occurred just before the start of sessions. Youth who were not randomized remained eligible for future sessions as long as they met the eligibility criteria at the time of randomization. Once youth were

[4] Potentially eligible youth were prioritized for academic support and recruitment based on the following: (1st) youth who had previously dropped out of school, (2nd) youth not currently on grade level, (3rd) youth who met other at-risk criteria.

[5] There are multiple versions of the consent form in order to ensure all eligible participants are able to be fully and adequately informed prior to consenting. All 17 year olds required parental consent and student assent. In addition to an English version Consent/Assent form there is a Spanish version of the Consent/Assent form. This was created to ensure potential participants who have Spanish-speaking parents are adequately informed. An Unaccompanied Minors Consent form was available for potential participants who were age 17, but did not have access to their legal guardian (e.g., parent in jail, unsafe for participant to contact parent). All youth age 18 and older were able to provide consent for themselves.

randomized into a cohort, they were no longer eligible to attend a future intervention regardless of their actual attendance. A detailed description of the steps that occurred prior to the random assignment process is presented in Appendix A.

The length of time between consent and randomization varied due to the rolling nature of collecting consents and the ability of potential participants to remain eligible for future interventions based on availability. Each available youth had an equal probability of being assigned to treatment or comparison within a particular cohort.

C. Data collection

Impact evaluation data were collected at four time points: baseline, 3 months, 6 months, and 12 months post-program. Youth who participated in the first three years of the evaluation contributed data at all four time points, while youth enrolled in the final year contributed data at only the first three time points. Data on program adherence, quality of implementation, and experiences of counterfactuals were collected at various time points to provide context for the impact findings.

1. Impact evaluation

The primary source of data for the impact evaluation was provided via self-report through an online survey participants completed at four time points (baseline, 3 months, 6 months, and 12 months post-program). Appendix B presents the timing for each data collection effort. In addition to the primary and secondary outcomes, the online survey collected information on sexual health knowledge, sexual health attitudes and beliefs, educational outcomes, and developmental assets and took approximately 30 minutes to complete. Baseline data collection for treatment participants concluded prior to start date of the intervention, while baseline data collection for comparison participants was extended until the final date of the intervention for the

same cohort of treatment participants.[6] Baseline surveys were primarily administered in a school setting, either individually or in a group. Less than 1% of youth were randomized and then dropped out of school prior to completing the baseline, thus requiring the baseline be administered outside of a school setting.

Follow-up surveys were conducted in different settings due to the variety of participant situations at the follow-up point.[7] If participants were still enrolled in school, they may have been surveyed in a school setting before or after classes or during class periods that were not core curriculum as determined by the graduation coach. If participants were no longer enrolled in school (e.g., graduated, dropped out, withdrawn), a member of the evaluation team attempted to schedule a time to administer the survey at a location and time that was convenient to the participant. All survey administrators had access to laptop computers and Wi-Fi hotspots, which allowed them to administer surveys at any time or location. Most surveys were administered at the participant's home, place of work, or a public location.

For the majority of surveys a member of the evaluation team was present to administer the survey; however, due to the transient nature of participants in this study an alternative procedure was created in order to allow participants to be surveyed who were unable to meet in person with a survey administrator (e.g., participant moved to another state or country, participant unable to meet during reasonable hours). A link to the survey along with the participant's random ID was sent to an e-mail address provided. Once the survey was verified as complete, an incentive and

[6] Due to limited resources and a condensed time schedule to complete the baselines it was necessary to allow for some flexibility in baseline data collection. It was expected that contamination would be minimized between groups since treatment participants were off campus for the duration of the three days of the intervention.

[7] Three-month follow-ups occurred 76-151 days from baseline. Six-month follow-ups occurred 152-243 days from baseline. Twelve-month follow-ups occurred 335-487 days from baseline. No one actively searched for participants 244-334 days from baseline; therefore, any surveys completed during this time frame were evaluated on a case by case basis to determine if it should be considered a 6-month or a 12-month follow-up.

thank you letter was delivered to the participant in person or by mail. Approximately 15% of surveys were administered in this manner.[8]

While the survey was offered primarily in an online format, there was also a paper and pencil version of the survey available for cases where technology failed or the participant was not comfortable with a computer and requested to take a paper survey. A member of the evaluation team manually entered paper surveys into the online survey system. Less than 1% of surveys were administered in this manner.

A $20 gift card to Walmart was provided to youth as an incentive to complete the surveys. Youth may have received a maximum of four gift cards (one for each data collection point) over the course of their participation in the project. Youth were encouraged to complete the entire survey, but were not required to complete it in full in order to receive the incentive, due to the sensitive nature of the questions. Some of the youth who could be reached for the 12-month survey were given the opportunity to attend a "Camp Thurman Fun Day" where youth were provided lunch and allowed to attend a ropes-based challenge course and use the equipment and activities. This was offered to all participants in the eligible cohorts regardless of their treatment status and was only an option following their final survey since Camp Thurman was part of the intervention; however, the activities from the intervention were not repeated during this day. Youth who completed all four surveys were entered into a raffle for an iPod to encourage youth to complete all of the data collection points.

[8] In order to be eligible to take the survey online without a survey administrator present the participant must have verbally agreed to take it online, verified their identity by providing their student ID, provided a valid e-mail address, and agreed to pick up their gift card in person or provide an address to mail the gift card to.

2. Implementation evaluation

The implementation study was conducted to assess program adherence, quality of implementation, experiences of counterfactuals, and context. See Appendix C for detail on the methods used to collect data for the implementation evaluation. Program adherence was assessed through number of sessions offered, average frequency of sessions, and the average number of activities attended by participants. It was further assessed by reporting on the percentage of activities fully completed as reported by program facilitators in their program observation logs at the completion of each intervention series. Lastly, the number and type of staff who delivered the program and percentage of staff attending staff development trainings was also assessed based on personnel records provided by the project director at the completion of the program.

Quality was assessed based on two measures: quality of staff-participant interactions, and quality of youth engagement with the program. A member of the evaluation team independently observed 75% of the offered program sessions and assessed both measures based on items from the Program Observation Form for TPP Grantees developed by HHS at the completion of each intervention series. Since multiple facilitators were administering the program in small groups, it was not possible to observe every facilitator for every interaction; therefore, these measures are an overall assessment of the entire intervention. It is possible that some facilitators would have ranked differently on interactions with participants had the reporting been done at the individual facilitator level; however, those data were not collected.

The experiences of the counterfactual were not formally assessed since there is no formal alternative program offered to counterfactual participants. A yearly interview with the coordinator of pregnancy related services, an expert on sexual health related services for adolescents enrolled in the school district, provided a summary of other potential sexual health related programming that youth might have participated in during the course of the study.

Context was assessed through periodic interviews with the pregnancy related services coordinator to determine if there were other TPP programming options offered to study participants or if there were any substantial external events affecting implementation. A summary of substantial unplanned adaptations to the curriculum was assessed based on a review of the fidelity logs generated following each intervention session and notes from fidelity meetings with facilitators.

D. Outcomes for impact analyses

Primary and secondary outcome measures are provided in Table III.1 and Table III.2, respectively. The data source for each of these measures is listed and comes directly from the survey or is derived from answers provided in the survey.

The primary research question is answered with a single-item dichotomous measure from the 3-month follow-up survey: "*In the past three months, have you had sexual intercourse without you or your partner using a condom?*"[9] This measure of risky sexual behavior, vaginal intercourse without a condom, captures the effect of offering *Crossroads* on the full analytic sample of youth, including youth who may or may not have sexually active[10] at baseline.

[9] Sexual intercourse was defined on the survey as the sexual act of a male putting his penis in a female's vagina, which will be referred to a vaginal intercourse for the duration of this report.

[10] Sexually active at baseline is defined as having ever engaged in any form of intercourse (i.e., vaginal intercourse, anal intercourse, or oral intercourse) at the baseline survey time point.

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

Outcome name	Description of outcome	Timing of measure relative to program
Vaginal intercourse without a condom	The variable is a yes/no measure of whether a person has had sexual intercourse in the past three month without using a condom. The measure is taken directly from the following item on the survey: • "In the past three months, have you had sexual intercourse without you or your partner using a condom?" The variable is constructed as a dummy variable where respondents who respond "No" were coded as 0 and responses of "Yes" were coded as 1. Youth who were not sexually active or did not engage in vaginal sexual intercourse were coded as 0.	3 months post-program

The secondary research questions are answered using the same variable from the primary research question at the 6-month and 12-month follow-up surveys. Two additional dichotomous items were constructed to assess risky sexual behavior with different types of sex assessed at all three follow-up data collection points. A final dichotomous item was constructed to assess pregnancy outcomes. A description of the secondary outcome measures is provided in Table III.2.

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

Outcome name	Description of outcome	Timing of measure relative to program
Vaginal intercourse without a condom	The variable is a yes/no measure of whether a person has had sexual intercourse in the past three month without using a condom. The measure is taken directly from the following item on the survey: • "In the past three months, have you had sexual intercourse without you or your partner using a condom?" The variable is constructed as a dummy variable where respondents who respond "No" were coded as 0 and responses of "Yes" were coded as 1. Youth who were not sexually active or did not engage in vaginal intercourse were coded as 0.	6 months post-program 12 months post-program

Outcome name	Description of outcome	Timing of measure relative to program
Anal intercourse without a condom	This variable is a dummy variable constructed from the numerical values taken directly from the survey • When you had anal sex in the past 3 months, how often were condoms (rubbers) used? The variable is constructed as a dummy variable where respondents who respond "Every time" were coded 0 and responses of "Never", "Sometimes", "Often", and "Almost every time" were coded as 1. Youth who were not sexually active or did not engage in anal intercourse were coded as 0.	3 months post-program 6 months post-program 12 months post-program
Oral intercourse without a condom	Dummy variable constructed from the numerical values taken directly from the survey for the following questions. • When you performed oral sex in the past 3 months, how often were condoms (rubbers) used? • When someone else performed oral sex on you in the past 3 months, how often were condoms (rubbers) used? The variable is constructed as a dummy variable where responses of "Every time" were coded as 0 and responses of "Never", "Sometimes", "Often", and "Almost every time" were coded as 1. Responses of "Never had intercourse" and "No oral sex in past 3 months" were coded as 0. These two variables were then combined into one variable. Responses of 0 on both newly constructed variables remained 0. Responses of 1 on both newly constructed variables remained 1. Responses of 1 on either newly constructed variable and 0 on the other newly constructed variable were coded to 1. Responses of 1 on either newly constructed variable and missing on the other newly constructed variable were coded as 1. Responses of 0 on either newly constructed variable and missing on the other newly constructed variable were coded as missing. Youth who had not engaged in any form of oral intercourse or were not sexually active were coded as 0.	3 months post-program 6 months post-program 12 months post-program
Ever been pregnant	The variable is a yes/no measure of whether a person has ever been or gotten someone pregnant. The measure is taken directly from the following item on the survey: • To the best of your knowledge, have you ever been pregnant or gotten someone pregnant? Responses of "No" were coded as 0. Responses of "Yes" were coded as 1. Youth who were not sexually active were coded as 0.	12 months post-program

E. Study sample

Appendix E represents the flow of sample members from the beginning of the study through the creation of analytic samples used for 3-month, 6-month, and 12-month follow-ups. At the

completion of the study, 1,896 youth were approached about participation in the study. Three hundred and one refused participation, 201 never returned a signed consent form, and 220 returned a consent form but were not randomized[11]. This resulted in 61.9% (n = 1,174) of eligible participants who were randomized at the individual level (n = 596 treatment and n = 578 comparison).

Table F.1 in Appendix F demonstrates the three analytic samples that were constructed to determine baseline equivalence. Eighty-two percent of randomized participants (n = 479 comparison and n = 478 treatment) responded to the baseline and 3-month follow-up survey, resulting in an overall attrition rate of 18.48% and a differential attrition rate of 2.67%. Seventy-seven percent of randomized participants (n = 429 comparison and n = 471 treatment) responded to the baseline and 6-month follow-up survey, resulting in an overall attrition rate of 23.34% and a differential attrition rate of 4.81%. Seventy percent of randomized participants eligible to take the long-term follow-up survey (n = 342 comparison and n = 377 treatment) responded to the baseline and 12-month follow-up survey, resulting in an overall attrition rate of 29.92% and a differential attrition rate of 4.37%.[12] Baseline equivalence was assessed on each analytic sample; however, not all of these participants responded to each question for the outcomes of interest. The final analysis for each outcome of interest had a slightly smaller sample size based on response rates to individual questions.[13]

[11] This includes youth who returned a consent form but were not available to attend any of the interventions or youth who consented but became ineligible (e.g., dropped out, graduated, moved) before randomization occurred.

[12] Fewer cohorts were eligible to provide 12-month follow-up data. Attrition rates are calculated based on the 1,026 participants who were eligible to participate in 12-month follow-up data collection, rather than the 1174 who were eligible for the 3-month and 6-month follow-up data points.

[13] Sample sizes included for each outcome of interest is included as a table notes for Table IV.1 – Table IV.4.

F. Baseline equivalence

Baseline equivalence was established for each analytic sample to assess the impact of attrition on the comparability of the treatment and comparison groups. A hierarchical linear model was used to account for the clustering of youth by cohort. Youth who provided baseline and follow-up data at each follow-up time point were assessed on the following variables: age, gender, race/ethnicity, grade, sexually active at baseline, and the outcomes of interest. Baseline equivalence for the youth responding to the 3-month follow-up is below; there were no statistically significant differences between the treatment and comparison groups at baseline. There were no statistically significant differences between treatment and comparison groups at 6-months and 12-months with the exception of the vaginal intercourse without a condom at 12-months (see Appendix G).

Table III.3: Summary statistics of key baseline measures for youth completing 3-month follow-up

Baseline measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison *p*-value of difference
Age	18.18 (.626)	18.10 (.633)	.08	.072
Gender (female)	49.2	48.4	0.8	.822
Race/ethnicity: White	10.3	11.7	-1.4	.406
Race/ethnicity: Black	33.9	28.2	5.7	.406
Race/ethnicity: Hispanic	34.5	36.7	-2.2	.406
Race/ethnicity: Other	1.7	2.3	-0.6	.406
Race/ethnicity: More than 1 race	19.7	21.1	-1.4	.406
Grade: 10th or lower	11.5	12.7	-1.2	.947
Grade: 11th grade	29.0	29.9	-0.9	.947
Grade: 12th grade	57.0	57.3	-0.3	.947
Grade: Other	1.3	1.3	0.0	.947
Sexually Active	79.7	78.1	1.6	.548
Vaginal intercourse without a condom	34.7	37.7	-3.0	.346

Baseline measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison *p*-value of difference
Anal intercourse without a condom	5.7	6.9	-1.2	.448
Oral intercourse without a condom	46.2	47.2	-1.0	.761
Sample size	478	479		

Note: The sample size reported here is for youth who provided a baseline and 3-month follow-up survey. Not all youth provided data on the four outcomes of interest and subsequently are based off of the following smaller sample sizes. Vaginal intercourse without a condom sample size was 476 treatment and 474 comparison. Anal intercourse without a condom sample size was 470 treatment and 468 comparison. Oral intercourse without a condom sample size was 466 treatment and 466 comparison.

G. Methods

1. Impact evaluation

The same analytic approach was used for both primary and secondary research questions. An intent-to-treat framework was used to assess the impact of the *Crossroads* program relative to the graduation coach program (business as usual counterfactual) on the primary research question regarding participants' engagement of vaginal intercourse without a condom three months after the end of treatment. The impact estimate is the regression-adjusted difference between the average outcomes of youth who were assigned to the intervention program and youth who were assigned to the comparison group. Impact estimates with *p*-values less than 0.05 are considered statistically significant and provide support that the differences in outcomes can be attributed to the *Crossroads* program. The primary analytic approach used a linear probability model with treatment group as the primary predictor along with student-level baseline characteristics as covariates (i.e., age, gender, race/ethnicity[14], grade, ever been sexually active at baseline, baseline measure of the outcome of interest, and cohort). The equation used for

[14] Race/Ethnicity was constructed as a collapsed dummy variable using the Race variable and Ethnicity variable. This resulted in five categories: Black/African American, White, Hispanic only, Other and More than one race where Black/African American was the omitted reference group.

estimating impacts is included in Appendix H. Subsequent sensitivity analyses were conducted without any covariates to further assess the robustness of the benchmark analyses. The results of those analyses are included in Appendix I.

Missing data were less than 10% and missing at random; therefore, complete case analysis was conducted and cases with missing data were ignored. Inconsistent responses were limited due to online data collection, which implemented skip logic to prevent inconsistent survey responses within the same survey.

2. Implementation evaluation

Descriptive statistics were used to assess program adherence, quality, counterfactual condition, and context for implementation evaluation. See Appendix D for methods used to address implementation research questions. To assess adherence to the program model, the following measures were assessed:

- Total number of sessions delivered
- Average frequency of sessions
- Average number of sample that attended 75% or more of the intervention[15]
- Average number of sample that did not receive any portion of intervention (no-shows)
- Mean number of activities attended by participants
- Average number of activities fully completed during each session
- Number and qualifications of staff who delivered the program
- Percentage of staff who attended staff development trainings.

[15] Seventy-five percent was based on attendance in hours. This was 15.75 out of 21 hours for the longer programming during cohorts 1 and 2, and 14 out of 18.75 hours for the shortened programming for cohorts 3 through 21.

Quality of program implementation was assessed as the percentage of observed interactions rated a 4 or higher on a scale of 1 to 5 on the following measures:

- Quality of rapport and communication with participants (1 = Poor and 5 = Excellent)
- Ability of staff to effectively address questions and concerns with participants (1 = Poor and 5 = Excellent)
- Participants levels of understanding (1 = Little understanding and 5 = Good understanding)
- Levels of participation for youth in group discussions and activities (1 = Little participation and 5 = Active participation)

No formal measures were used to assess the counterfactual experience for youth. This is a substantial limitation of the implementation evaluation and resulted in an unclear picture of the counterfactual experience for participants. A general description of other services youth may have participated in is provided based on interviews from district staff.

No formal measures to assess context were collected; rather, it was assessed based on annual reports where substantial external events affecting implementation were discussed. A summary of these events and any significant unplanned adaptations to the curriculum is provided.

IV. Study findings

A. Implementation study findings

A total of 21 sessions were offered across three and a half school years. An average of one session per month was held during the months programming was offered. Almost 71% of youth attended 75% or more of the intervention, while 13.6% of the sample did not receive any portion of intervention. All 16 intended activities were offered per session resulting in an average of 13.06 (*s.d.* 6.06) activities attended by participants

Fidelity was high with 96% of activities observed being fully completed as intended. A total of four qualified staff, comprised of one certified teacher, one licensed professional counselor, and two licensed masters level social workers, delivered the program to participants for the duration of the project. Three of the four staff attended 2-3 staff development trainings per academic year to assist staff to better work with at risk youth.

The program exhibited high quality on both measures assessing the quality of staff-participant interactions. All of the observations were rated at a 4 or higher, demonstrating that participants had a "good understanding" of the material. Nearly all (96.2%) of the observations were rated at a 4 or higher, demonstrating "active participation" among youth during discussions and activities. Quality of youth engagement was also high. All of the observations were rated at a 4 or higher, demonstrating "excellent" levels of rapport and communication between staff and youth. All were rated at a score of 5 ("excellent") regarding the ability of the facilitator to effectively address concerns and questions from youth.

The counterfactual experience for youth was not formally assessed. Both youth in treatment and comparison had access to graduation coaches who provided ongoing academic support and Drop Out Prevention services. None of these services were aimed at addressing sexual health topics. No alternative comprehensive sexual health programs were offered to youth within the school district or local community. Since fall of 2011, younger students have had an opportunity to participate in *Promoting Health Among Teens*, an abstinence only based curriculum; however, none of these students would have been eligible to participate in *Crossroads* based on their age. Still, it is possible that siblings, relatives, or friends of *Crossroads* participants may have participated in this program and subsequently communicated information regarding the content to study participants. Youth may have received some content regarding sexual health in health

classes offered by the school district.[16] Youth who were pregnant or parenting may have received additional programming and support from a parenting education program.[17] Lastly, youth may have had access to community-based services such as local health clinics. Interviews with program staff and an online search found one formal program available to youth in Arlington that is directed at pregnancy prevention. This program targets females age 6-18 and provides a multitude of services outside of preventing adolescent pregnancy. It is unknown if any of the youth in the study participated in this program.

Interviews with program staff and the Pregnancy Related Services coordinator did not reveal any substantial external events affecting implementation on an ongoing basis. One of the cohorts did experience severe weather preventing youth from participating in the outdoor experiential activities; however, the staff was able to modify the programming in order to continue to implement the curriculum and cover all core components. Only one substantial unplanned adaptation occurred during the course of the project. The original length of the programs included 21 hours of curriculum across three days. It was necessary to shorten the length of the individual days in order to ensure youth were returned to campus by the end of the school day. Starting with the third cohort, the length of the program was reduced to 18.75 hours across the three days. To accommodate this reduction in time, activities that were redundant or non-essential were removed from the curriculum. No core components were removed.

[16] No formal assessment of content delivered in health classes was performed. An overview of the Texas Essential Knowledge and Skills, or state standards for Texas public schools, for Health Education in high schools can be found in Texas Education Code (Texas Essential Knowledge and Skills for Health Education, T.E.C §28.002, 1998).

[17] The Pregnancy, Education, and Parenting program offered through the school district, provides comprehensive services to youth who are pregnant or parenting. This service provides access to a school nurse, along with individual and group counseling to address the needs of young parents.

B. Impact study findings

The impact study findings are presented by primary and secondary research questions in tables IV.1 through IV.4.

1. Primary research question – Vaginal intercourse without a condom at 3 month follow-up

Analysis for the primary question regarding condom use during vaginal intercourse was conducted using linear probability modeling that included the covariates for gender, age, race/ethnicity, grade, ever been sexually active at baseline, cohort[18] and condom use during vaginal intercourse at baseline. At the 3-month follow-up, within the comparison group, 36.7% responded that they did not use a condom during vaginal intercourse, compared to 35.0% of the treatment group with a similar response. Results from the regression model indicated that there was no statistically significant difference between the two groups in condom use during vaginal intercourse at 3-month follow-up (p = .552).

Table IV.1. Post-intervention estimated effect using data from 3-month follow-up to address the primary research question

Outcome measure	Treatment %	Comparison %	Treatment effect (p-value of difference)
Sexual Intercourse without a Condom	35.0	36.7	-1.7 (.552)

Source: Follow-up surveys administered 3 months after the program.

Notes: See Table III.1 for a more detailed description of each measure and section III for a description of the impact estimation methods. The sample used in this analysis included youth who responded to both the baseline and 3-month follow-up (treatment n = 476 and comparison n = 474).

2a. Secondary research questions – Anal and oral intercourse without a condom at 3-month follow-up

With the inclusion of covariates, results from the regression model at the 3-month follow-up, 7.0% of the comparison group and 6.0% of the treatment group indicated they had engaged in

[18] Cohort was assessed using a series of dummy variables (x-1).

anal intercourse without using a condom; 41.0% of the comparison group and 42.7% of the treatment group indicated they had engaged in oral intercourse without using a condom. There was no significant difference between the two groups in condom use during anal intercourse (p = .573) or in condom use during oral intercourse (p =.581).

Table IV.2 Post-intervention estimated effects using data from 3-month follow-up to address the secondary research questions

Outcome measure	Treatment %	Comparison %	Treatment effect (p-value of difference)
Anal Intercourse without a Condom	6.0	7.0	-1.0 (.573)
Oral Intercourse without a Condom	42.7	41.0	1.7 (.581)

Source: Follow-up surveys administered 3 months after the program.

Notes: Each outcome measure was assessed based on a separate analytic sample. See Table III.3 for a more detailed description of each measure and section III for a description of the impact estimation methods. Anal intercourse without a condom sample size was 470 treatment and 468 comparison. Oral intercourse without a condom sample size was 466 treatment and 466 comparison.

2b. Secondary research questions – Vaginal, Anal, and Oral intercourse without a condom at 6-month follow-up

At the six-month follow-up, 40.4% of the comparison group responded that they did not use a condom during vaginal intercourse, compared to 31.7% of the treatment group. After conducting the linear probability model, statistically significant differences between the comparison and treatment groups $\left(p\langle.05\right)$ were found, indicating that membership in the treatment group at the six-month follow-up may have contributed to a decrease in risky sexual behavior. These findings were not replicated across other forms of intercourse. Neither of the models for anal or oral intercourse without a condom at the 6-month follow-up demonstrated group membership as a significant predictor (p = .866 and p = .477, respectively) indicating no intervention impacts on these two types of risky sexual behavior.

Table IV.3 Post-intervention estimated effects using data from 6-month follow-up to address the secondary research questions

Outcome measure	Treatment %	Comparison %	Treatment effect (*p*-value of difference)
Vaginal Intercourse without a Condom	31.7	40.4	-8.7 (.006)
Anal Intercourse without a Condom	7.4	7.1	0.3 (.866)
Oral Intercourse without a Condom	42.7	40.4	2.3 (.477)

Source: Follow-up surveys administered 6 months after the program.

Notes: Each outcome measure was assessed based on a separate analytic sample. See Table III.2 for a more detailed description of each measure and section III for a description of the impact estimation methods. Vaginal intercourse without a condom sample size was 465 treatment and 419 comparison. Anal intercourse without a condom sample size was 465 treatment and 421 comparison. Oral intercourse without a condom sample size was 453 treatment and 416 comparison.

2c. Secondary research questions – Vaginal, Anal, and Oral intercourse without a condom and Pregnancy at 12-month follow-up

None of the regression models across sexual behavior outcomes at the 12-month follow up indicated that group membership was a significant predictor of the outcome of interest, indicting that the intervention did not significantly impact risky sexual behavior at the long-term follow-up.

The pregnancy variable was analyzed at only the 12-month follow-up and sought to determine if the participant had ever been pregnant or gotten someone pregnant. A larger percentage of the treatment group reported pregnancies (24.7%) compared to the comparison group (22.4%). However, the regression model demonstrated that there was no statistically significant difference between the two groups ($p = .378$).

Table IV.4 Post-intervention estimated effects using data from 12-month follow-up to address the secondary research questions

Outcome measure	Treatment %	Comparison %	Treatment effect (*p*-value of difference)
Vaginal Intercourse without a Condom	40.2	42.9	-2.7 (.419)
Anal Intercourse without a Condom	8.7	9.3	-0.6 (.776)
Oral Intercourse without a Condom	43.3	43.4	-0.1 (.972)
Pregnancy	24.7	22.4	2.3 (.378)

Source: Follow-up surveys administered 12 months after the program.

Notes: Each outcome measure was assessed based on a separate analytic sample. See Table III.2 for a more detailed description of each measure and section III for a description of the impact estimation methods. Vaginal intercourse without a condom sample size was 373 treatment and 335 comparison. Anal intercourse without a condom sample size was 357 treatment and 331 comparison. Oral intercourse without a condom sample size was 357 treatment and 327 comparison. Pregnancy sample size was 351 treatment and 318 comparison.

V. Conclusion

The results of this study indicate that *Crossroads*, an adapted version of the *Be Proud! Be Responsible!* curriculum, was not effective at reducing risky sexual behavior among older adolescents at the short-term follow-up. The curriculum was selected based on its proven effectiveness with reducing risky sexual behavior through increased condom use and reduction in the frequency of unprotected sex (Jemmott, Jemmott, & Fong, 1998; Jemmott, Jemmott, Braverman, & Fong, 2005). While these studies found positive outcomes at 12-months post-programming, it was anticipated that outcomes would be most prevalent at the short-term follow-up (3 months) for the older adolescent population; however, findings determined this was not the case. Rather, the only measurable impact of *Crossroads* occurred at the six-month follow-up, which may be attributed to other factors not previously considered. Specifically, evidence from interviews with youth indicated that changes in youth's sexual behaviors (e.g., increased condom use) and personal successes (e.g., graduating high school, accessing sexual health resources) were a direct result of their participation in the program and was closely related to the

relationship developed with their facilitator (Slater & Mitschke, 2015). The decision to have facilitators make monthly follow-up phone calls to youth to maintain contact for follow-ups and address concerns related to the intervention may have had an unanticipated effect impacting this outcome. As facilitators continue to make follow-up contacts, youth begin to develop a stronger relationship with them, establish higher levels of trust, and subsequently become more comfortable and willing to utilize the skills and knowledge provided during the intervention. While it would be expected that this would carry over to the 12-month follow-up as well, it did not. This is likely due to the changing school and life circumstances for youth at this time point. Since this study focused on older adolescents, many had graduated or moved on to a different phase of their life, and subsequently transitioned out of adolescence by the 12-month follow-up point. Anecdotal evidence from youth indicated that many of the participants in the program had transitioned into long term committed relationships, including marriage, resulting in a shift in condom use decisions, which may have impacted the long-term findings. These findings are consistent with another study of racially diverse youth that failed to find significant program effects at 12-months post programming (Borawski, Traple, Adams-Tufts, Hayman, Goodwin, & Lovegreen, 2009).

This study failed in the attempt to assess the effectiveness of *Crossroads* on risky anal and oral sexual behaviors. Anal intercourse, in particular, demonstrated a small proportion of the youth who reported to be engaged in this type of intercourse. The incorporation of these questions were intended to be more inclusive of gay youth; however, the assessment of whether these were the youth captured with this data is limited since no variables were collected regarding relationship status or types of relationships that were involved during intercourse. Without this information it is difficult to understand if the low numbers of youth reporting

engaging in anal intercourse was due to social stigma of reporting engaging in this type of intercourse for either gay or heterosexual youth, or if this reflects the subpopulation of gay youth who are less likely to be in a stable placement and thus more difficult to locate at follow up time points (Durso & Gates, 2012). In contrast, a much larger sample reported engaging in oral sex without condoms. It did appear that the majority of people engaging in oral intercourse did not use a condom regardless of group status, which may be more reflective of societal norms rather than the ineffectiveness of the program.

While the program overall was implemented with high fidelity and quality, the lack of significant positive findings is most likely due to age of participants. This study is unique in that it attempts to address the needs of older adolescents, and while *Be Proud! Be Responsible!* was much more effective with youth as young as age 12 (Jemmott, Jemmott, & Fong, 1998), it did not do the same with older youth. Older adolescents have likely established sexual norms by the time they are age 17, making it more difficult to change overall behaviors. The majority (80%) of the overall sample in this study reported they were already sexually active at baseline. This in conjunction with the fact that almost a quarter of this sample of youth reported having already experienced a pregnancy clearly represents the need to address sexual health with older adolescents; however, the *Crossroads* program and curriculum may be more suitable and effective for a younger population who has not already engaged in sexual activity, but for whom the possibility is a plausible choice.

The *Crossroads* program was implemented effectively as intended; however, fundamental challenges surrounding the structure of the program that should be addressed before implementing a similar program within a school setting in the future. Targeting academically at-risk youth for an intervention that required youth to miss three days of schools was an ongoing

issue for the duration of the project. Teachers, staff, parents, and youth presented understandable resistance, which contributed to low recruitment and retention numbers, in spite of numerous attempts to address these concerns.[19] Reducing the program to one day and offering it on days when youth are not required to be in school may be a more effective strategy to overcome this challenge, though it would likely prove challenging to deliver all the major content in just 6-8 hours.

Understanding the experience of youth in the counterfactual condition should be addressed in the future. As the study progressed, many youth were recruited to the study by their friends and relatives who participated in *Crossroads*, which speaks to the positive experiences youth had with the program. Over time, siblings, boyfriends and girlfriends, and friends signed up for the study together. As a result, contamination likely occurred since youth assigned to both conditions were enrolled at each campus and in varying types of relationships with one another. Addressing this type of contamination would be critical for future studies within a school setting.

The presence of impacts for *Crossroads* at six months post-intervention on one outcome, vaginal intercourse without a condom, is promising and provides the foundation for future studies. The absence of these impacts at the short and long term follow-ups further suggest the need to develop programming for older adolescents in order to adequately and effectively meet the needs of this overlooked population.

[19] Attempts to address these concerns included scheduling the program on days that did not conflict with district-wide testing, the beginning or end of the six-week period, and school related events. Shortening the length of the day and providing transportation for youth who needed to return back to their home campus during the programming was also offered. Graduation coaches were also proactive in communicating with teachers and helping youth by getting missed assignments and scheduling make-up work prior to the youth attending the intervention.

VI. References

The Annie E. Casey Foundation. *The 2008 KIDS COUNT data book*. Baltimore: The Annie E. Casey Foundation, 2008. Accessed November 1, 2015. http://www.aecf.org/resources/the-2008-kids-count-data-book/.

Arlington Independent School District. (2011) *Pregnancy Related Services Database System.* [Data file]. Available from http://aisd.net.

Arlington Independent School District. (2009) *Academic Excellence Indicator System.* [Data file]. Available from http://aisd.net.

Arlington Independent School District. (2008) *Academic Excellence Indicator System.* [Data file]. Available from http://aisd.net.

Borawski, E.A., Traple, E.S., Adams-Tufts, K., Hayman, L.L., Goodwin, M.A. & Lovegreen, L.D. (2009). Taking Be Proud! Be Responsible! to the suburbs: A replication study. *Perspectives on Sexual and Reproductive Health, 41*(1), 12-22.

Crosby, R.A., Diclemente, R.J., Wingood, G.M., Slazar, L.F., Rose, E., & Sales, J.M. (2007). The protective value of school enrollment against sexually transmitted disease: A study of high-risk African American adolescent females. *Sexually Transmitted Infections, 83*(3), 223-227.

Durso, L.E., & Gates, G.J. (2012). *Serving our youth: Findings from a national survey of service providers working with lesbian, gay, bisexual and transgender youth who are homeless or at risk of becoming homeless.* Los Angeles, CA: The Williams Institute with True Colors Fund and The Palette Fund.

Frisco, M.L. (2008). Adolescents' sexual behavior and academic attainment. *Sociology of Education, 81*(3), 284-311.

Jemmott, L. S., & Jemmott, J. B. III (1992). Increasing condom-use intentions among sexually active inner-city adolescent women: Effects of an AIDS prevention program. *Nursing Research, 41,* 273-278.

Jemmott, J.B. III, Jemmott, L.S., Braverman, P.K., & Fong G.T. (2005). HIV/STD risk reduction interventions for African American and Latino adolescent girls at an adolescent medicine clinic: A randomized controlled trial. *Archives of Pediatric Adolescent Medicine, 159,* 440-449.

Jemmott, J.B., Jemmott, L.S., & Fong, G.T. (1998). Abstinence and safer sex HIV risk-reduction interventions for African American adolescents. *Journal of the American Medical Association, 279*(19)*,* 1529-1536.

Jemmott, J. B., Jemmott, L. S., & Fong, G. T. (1992). Reductions in HIV risk-associated behaviors among black male adolescents: Effects of an AIDS prevention initiative. *American Journal of Public Health*, *82*(3)*,* 372-377.

Jemmott J.B., Jemmott, L.S., Fong, G.T., & McCaffree, K. (1999). Reducing HIV risk-associated sexual behavior among African American adolescents: Testing the generality of intervention effects. *American Journal of Community Psychology*, *27*(2)*,* 161-187.

Jemmott, L.S., Jemmott, J.B. III, & McCaffree, K.A. (1996). *Be Proud! Be Responsible! Strategies to empower youth to reduce their risk for AIDS, Curriculum Manual*. New York: Select Media.

National Campaign to Prevent Teen and Unplanned Pregnancy. (2012). Why it matters: Teen childbearing, education, and economic wellbeing. Retrieved October 2014 from http://thenationalcampaign.org/sites/default/files/resource-primary-download/childbearingeducation-economicwellbeing.pdf

Perper, K., Peterson, K., & Manlove, J. (2010). Diploma Attachment Among Teen Mothers. Washington, DC: Child Trends. Retrieved October 2014, from http://www.childtrends.org/Files//Child_Trends-2010_01_22_FS_DiplomaAttainment.pdf

Slater, H.M., & Mitschke, D.B. (November, 2015). *Capturing Hidden Program Effects: Closing Out Evaluation of Program with Qualitative Interviews.* Poster presented at the American Evaluation Association Annual Conference in Chicago, IL.

Tarrant County Health Department. (2009). Sexually transmitted diseases in Tarrant county, 2001-2008.

Texas Essential Knowledge and Skills for Health Education. Texas Education Code §28.002.

Appendix A: Description of Random Assignment Process

A detailed description of the steps that occurred prior to the random assignment process is presented in below.

Step 1: Identify youth that met eligibility criteria. This occurred approximately two weeks prior to the intervention. The evaluation team supplied a list of all eligible youth from across the district that met the eligibility criteria to graduation coaches. The goal for the project was to have a minimum of 100 eligible youth who met the inclusion criteria prior to verifying availability (Step 2); however, it was often extremely difficult to meet this minimum number.

Step 2: Verify availability. This occurred approximately two weeks prior to intervention. Graduation coaches had 3-5 days to make contact with eligible youth and verify their availability and willingness to participate in the study and the intervention should they be selected for the treatment group. A list of eligible youth with an updated availability status (i.e., Available, Not Available, No Longer Eligible, Unable to Contact) was then provided to the evaluation team.

Step 3: Randomization. Approximately one week prior to intervention series a block/cohort of 60 youth who met the eligibility criteria and verified availability were randomly assigned to condition (30 Treatment and 30 Comparison) prior to each intervention series by the evaluation team.

Each available youth was assigned a random computer generated number. Randomization occurred by sorting available youth by random number and assigning every other available youth to treatment or comparison. If there were more than 60 available youth, youth were assigned a random computer generated number, sorted, and only the first 60 were randomized. The

remaining available youth who were not randomized remain eligible for future interventions. In situations where there were not 60 available youth at the time of randomization, graduation coaches continued to recruit and check availability for eligible youth they were unable to contact. As additional youth were identified as available a member of the evaluation team randomly assigned them to treatment or comparison until the cohort was filled. One exception to this occurred when 59 youth had been assigned and there was only one slot left. In this situation, the final slot was left empty to avoid bias in the assignment of a participant to a predetermined condition. Youth that were randomized as they confirmed they were available on an individual basis were assigned to condition by pulling out of a hat which group (treatment or comparison) they were assigned to. If a youth confirmed availability after the 60 slots were filled they remained eligible to be randomized in a future cohort and were checked for availability for future intervention. Once youth became randomized into a cohort they were no longer eligible to attend a future intervention regardless of their actual attendance.

Appendix B: Data collection efforts

Table B.1 Data collection efforts used in the impact analysis of *Crossroads* and timing

Data collection effort	Cohort 1	Cohort 2	Cohort 3	Cohort 4	Cohort 5	Cohort 6	Cohort 7	Cohort 8
Start date of programming	11/01/11	11/29/11	02/07/12	03/27/12	04/03/12	10/02/12	10/30/12	11/27/12
Baseline survey	10/25/11 – 11/3/11	11/22/11 – 12/01/11	01/31/12 – 02/09/12	03/20/12 – 03/29/12	03/27/12 – 04/05/12	09/25/12 – 10/04/12	10/23/12 – 11/01/12	11/20/12 – 11/29/12
3-month follow-up	01/16/12 – 03/31/12	02/13/12 – 04/28/12	04/23/12 – 07/07/12	06/11/12 – 08/25/12	06/18/12 – 09/01/12	12/17/12 - 03/02/13	01/14/13 – 03/30/13	02/11/13 – 04/27/13
6-month follow-up	04/01/12 – 07/01/12	04/29/12 – 07/29/12	07/08/12 – 10/07/12	08/26/12 - 11/25/12	09/02/12 – 12/02/12	03/03/13 – 06/02/13	03/31/13 - 06/30/13	04/28/13 – 07/28/13
12-month follow-up	10/01/12 – 03/02/13	10/29/12 – 03/30/13	01/07/13 – 06/08/13	02/25/13 – 07/27/13	03/04/13 – 08/03/13	09/02/13 – 02/01/14	09/30/13 – 03/01/14	10/28/13 – 03/29/14
Data collection effort	**Cohort 9**	**Cohort 10**	**Cohort 11**	**Cohort 12**	**Cohort 13**	**Cohort 14**	**Cohort 15**	**Cohort 16**
Start date of programming	01/22/13	02/12/13	03/26/13	04/20/13	10/08/13	11/19/13	01/28/14	02/25/14
Baseline survey	01/15/13 – 01/24/13	02/05/13 – 02/14/13	03/19/13 – 03/28/13	04/13/13 – 04/22/13	10/01/13 – 10/10/13	11/12/13 – 11/21/13	01/21/14 – 01/30/14	02/18/14 – 02/27/14
3-month follow-up	04/08/13 – 06/22/13	04/29/13 – 07/13/13	06/10/13 – 08/24/13	07/05/13 – 09/18/13	12/23/13 - 03/08/14	02/03/14 – 04/19/14	04/14/14 – 06/28/14	05/12/14 – 07/26/14
6-month follow-up	06/23/13 – 09/22/13	07/14/13 – 10/13/13	08/25/13 - 11/24/13	09/19/13 – 12/19/13	03/09/14 – 06/08/14	04/20/14 – 07/20/14	06/29/14 – 09/28/14	07/27/14 – 10/26/14
12-month follow-up	12/23/13 – 05/24/14	01/13/14 – 06/14/14	02/24/14 – 07/26/14	03/21/14 – 08/20/14	09/08/14 – 02/07/15	10/20/14 – 03/21/15	12/29/14 – 05/30/15	01/26/15 – 06/27/15

Table B.1 Data collection efforts used in the impact analysis of *Crossroads* and timing (cont.)

Data collection effort	Cohort 17	Cohort 18	Cohort 19	Cohort 20	Cohort 21
Start date of programming	03/25/14	04/29/14	10/07/14	10/28/14	11/18/14
Baseline survey	03/18/14 – 03/27/14	04/22/14 – 05/01/14	09/30/14 – 10/09/14	10/21/14 – 10/30/14	11/11/14 – 11/20/14
3-month follow-up	06/09/14 – 08/23/14	07/14/14 – 09/27/14	12/22/14 – 03/07/15	01/12/15 - 03/28/15	02/02/15 – 04/18/15
6-month follow-up	08/24/14 – 11/23/14	09/28/14 – 12/28/14	03/08/15 – 06/07/15	03/29/15 – 06/28/15	04/19/15 – 07/19/15
12-month follow-up	02/23/15 – 07/25/15	03/30/15 – 07/31/15*	N/A	N/A	N/A

N/A = Not Applicable

Appendix C: Implementation evaluation data collection

Table C.1. Data used to address implementation research questions

Implementation element	Types of data used to assess whether the element of the intervention was implemented as intended	Frequency/sampling of data collection	Party responsible for data collection
Adherence: How often were sessions offered? How many were offered?	All sessions delivered are captured in the *Crossroads* database. Number of cohort/sessions delivered Average frequency of sessions	All sessions delivered are captured in the *Crossroads* database.	Program staff
Adherence: What and how much was received?	All daily attendance records are captured in the Crossroads database.	Student attendance at all sessions is captured in the *Crossroads* database	Program staff
Adherence: What content was delivered to youth?	Number of activities not done, partially completed, and fully completed are captured on fidelity observation logs.	Observations occurred at a minimum of 75% of the offered interventions.	Evaluation staff
Adherence: Who delivered material to youth?	List of staff members hired and trained to implement program Background qualifications of staff members from staff applications List of training logs # and type of staff delivering the program to participants Program staff job requirements	Data on all staff members are available to program staff.	Project Director
Quality: Quality of staff-participant interactions	Observations of interaction quality using Program Observation Form for TPP Grantees developed by HHS	Random sample of 10% of program sessions observed	Evaluation staff
Quality: Quality of youth engagement with program	Observations of interaction quality using Program Observation Form for TPP Grantees developed by HHS	Random sample of 10% of program sessions observed	Evaluation staff
Counterfactual: Experiences of comparison condition	Interview with Pregnancy Related Services Coordinator	Once per year	Evaluation staff

Implementation element	Types of data used to assess whether the element of the intervention was implemented as intended	Frequency/sampling of data collection	Party responsible for data collection
Context: Other TPP programming available or offered to study participants (both intervention and comparison)	Interview with Pregnancy Related Services Coordinator	Once per year	Evaluation staff
Context: External events affecting implementation	Interview with Pregnancy Related Services Coordinator	Once per year	Evaluation staff
Context: Substantial unplanned adaptation(s)	Fidelity summary reports Adaptation requests, work plan, 6-month progress reports, annual progress reports	Generated after every session Annually	Program staff, project director, evaluation staff

Appendix D: Implementation evaluation methods

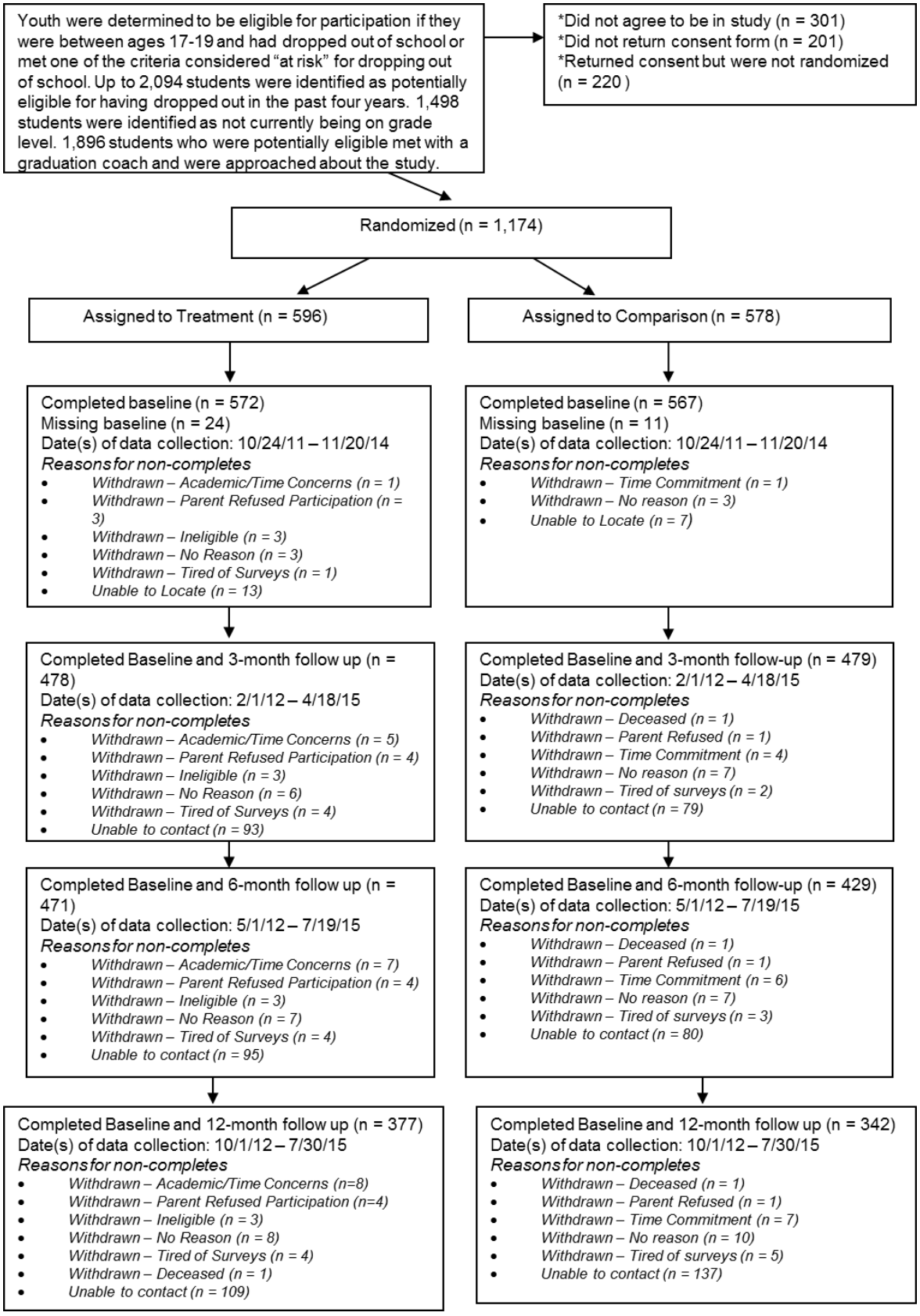
Table D.1. Methods used to address implementation research questions

Implementation element	Methods used to address each implementation element
Adherence: How often were sessions offered? How many were offered?	The total number of sessions is the sum of the sessions captured in the *Crossroads* database. Average frequency is calculated as the total number of sessions divided by the total number of months when programming was offered.
Adherence: What and how much was received?	Percentage of the sample that attended 75% or more of the intervention is calculated as the percentage of youth that attended 75% or more of the intervention they were assigned to. 75% of the intervention was calculated based on the total number of hours offered for a particular cohort. (Note: the shortened length of time for Cohorts 1 and 2 was taken into account when calculating percentages. This was calculated as 15.75 hours out of 21 hours for Cohorts 1 and 2 and 14 hours out of 18.75 hours for all remaining cohorts.)
Adherence: What content was delivered to youth?	The percentage of activities fully completed was a combination of all topics observed by an independent observer. An observer was present at 75% of all offered interventions. All observed activities were marked as not done, partially completed, or fully completed. The total number of topics marked "fully completed" divided by the total number of topics observed was used to calculate percentage of activities fully completed. (Note: a limitation to this measure is that the observer was not able to observer every small-group facilitator at the same time. As a result, some facilitators may have covered the topics completely while others did not. To attempt to mitigate this the observer rotated between facilitators during different interventions in order to observe all facilitators at multiple time points.)
Adherence: Who delivered material to youth?	Total number and type of staff delivering the program is a simple count of staff members implementing the program, taking into account their characteristics. Program staff job requirements includes a summary of the job requirements used for the job posting. Percentage of staff attending staff development training is calculated as the number of staff attending training divided by the total number of staff attending two or more trainings per academic year.
Quality: Quality of staff-participant interactions	One indicator of staff-participant interactions is calculated as the percentage of observed interactions in which the independent evaluator scored the intervention quality of rapport and communication with participants as 4 or higher on a scale of 1 to 5, where 1 = Poor (e.g., doesn't remember names, doesn't connect with participants, acts distant or unfriendly), 3 = Average, and 5 = Excellent (e.g., gets participants talking and excited, friendly, seems to understand the group and its needs). A second indicator of staff-participant interactions is calculated as the percentage of observed interactions in which the independent evaluator scored the ability of staff to effectively address questions and concerns with participants as 4 or higher on a scale of 1 to 5, where 1 = Poor (e.g., engages in "power struggles", gives inaccurate information, doesn't direct participants to additional resources), 3 = Average, and 5 = Excellent (e.g., answers questions with factual information, able to direct to appropriate information if doesn't know the answer).A random sample of 10% of observations was used to calculate these scores. (Note: Independent observers provided an overall assessment of these measures based on their observations throughout each day of the intervention. It is possible that these scores would have more variability if these measures were assessed for each activity and/or facilitator rather than an overall general score.)

Implementation element	Methods used to address each implementation element
Quality: Quality of youth engagement with program	One indicator of youth engagement with the program is calculated as the percentage of observed interactions in which the independent evaluator scored the participants levels of understanding as 4 or higher on a scale of 1 to 5, 1 = Little understanding (or roughly less than 25% seemed to understand), 3 = Some understanding (about half seemed to understand), and 5 = Good understanding (75 – 100% seemed to understand). A second indicator of youth engagement is calculated as the percentage of observed interactions in which the independent evaluator scored the levels of participation for youth in group discussions and activities as 4 or higher on a scale of 1 to 5, where 1 = Little participation (or roughly less than 25% participated), 3 = Some participation (about half participated), and 5 = Active participation (75 – 100% participated). A random sample of 10% of observations was used to calculate these scores. (Note: There are limitations to this data due to the data being collected by an independent observer and no formal measures being taken from the participants directly. Independent observers provided an overall assessment of these measures based on their observations throughout each day of the intervention. It is possible that these scores would have more variability if these measures were assessed for each activity and/or facilitator rather than an overall general score.)
Counterfactual: Experiences of counterfactual condition	The counterfactual experience is presented as a summary statement based on individual interviews conducted with the Pregnancy Related Services Coordinator. (Note: A major limitation of this measure is that no formal measure was collected directly from participants.)
Context: Other TPP programming available or offered to study participants (both intervention and counterfactual)	Other TPP programming available to study participants is presented as a summary statement based on individual interviews conducted with the Pregnancy Related Services Coordinator.
Context: External events affecting implementation	A summary statement of external events affecting implementation based on individual interviews conducted with the Pregnancy Related Services Coordinator is included.
Context: Substantial unplanned adaptation(s)	The unplanned change in program delivery setting is indicated. The resulting change in time allocated for facilitation of sessions is also described.

TPP = Teen Pregnancy Prevention

Appendix E: Intake Data Flow Chart - CONSORT Diagram for Youth



Appendix F: Study sample

Table F.1. Youth sample sizes by intervention status

Number of youth	Time Period	Total sample size	Intervention sample size	Comparison sample size	Total response rate	Intervention response rate	Comparison response rate
Assigned to condition		1,174	596	578			
Eligible for 12-Month Survey		1,026	522	504			
Contributed a baseline survey		1,139	572	567	.970	.960	.981
Contributed to baseline and follow-up survey							
	3 months post-programming	957	478	479	.815	.802	.829
	6 months post-programming	900	471	429	.767	.790	.742
	12 months post-programming	719	377	342	.701	.722	.679

All participants were eligible for inclusion in the 3 and 6-month follow-up sample, while a smaller number of participants was eligible for inclusion in the 12-month follow-up sample. Response rates for the 12-month post-programming were calculated using the smaller sample that was eligible for the 12-month survey.
N/A = Not Applicable

Appendix G: Baseline Equivalence Tables

Table G.1: Summary statistics of key baseline measures for youth completing 6-month follow-up survey

Baseline measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison *p*-value of difference
Age	18.17 (.626)	18.10 (.633)	0.07	.072
Gender (female)	49.3	50.8	-1.5	.640
Race/ethnicity: White	10.2	11.2	-1.0	.777
Race/ethnicity: Black	33.8	29.8	4.0	.777
Race/ethnicity: Hispanic	33.5	36.1	-2.6	.777
Race/ethnicity: Other	2.3	2.6	-0.3	.777
Race/ethnicity: More than 1 race	20.2	20.3	-.01	.777
Grade: 10th or lower	12.1	12.1	0.0	.990
Grade: 11th grade	31.6	32.6	-1.0	.990
Grade: 12th grade	54.8	53.8	1.0	.990
Grade: Other	1.5	1.4	0.1	.990
Sexually Active	79.1	77.9	1.2	.649
Vaginal Intercourse without a Condom	34.0	38.6	-4.6	.153
Anal intercourse without a Condom	6.0	6.3	-0.3	.859
Oral Intercourse without a Condom	44.4	45.7	-1.3	.703
Sample size	471	429		

Note: The sample size reported here is for youth who provided a baseline and 6-month follow-up survey. Not all youth provided data on the four outcomes of interest and subsequently are based off of the following smaller sample sizes. Vaginal intercourse without a condom sample size was 465 treatment and 419 comparison. Anal intercourse without a condom sample size was 465 treatment and 421 comparison. Oral intercourse without a condom sample size was 453 treatment and 416 comparison.

Table G.2: Summary statistics of key baseline measures for youth completing 12-month follow-up survey

Baseline measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison *p*-value of difference
Age	18.11 (.625)	18.18 (.619)	-0.07	.106
Gender (female)	49.1	49.4	-0.3	.927
Race/ethnicity: White	11.1	13.2	-2.1	.669
Race/ethnicity: Black	32.9	28.7	4.2	.669
Race/ethnicity: Hispanic	34.5	37.7	-3.2	.669
Race/ethnicity: Other	2.4	2.6	-0.2	.669
Race/ethnicity: More than 1 race	19.1	17.8	1.3	.669
Grade: 10th or lower	13.8	13.5	0.3	.754
Grade: 11th grade	32.6	34.8	-2.2	.754
Grade: 12th grade	52.8	50.3	2.5	.754
Grade: Other	0.8	1.5	-0.7	.754
Sexually Active	77.3	76,9	-0.4	.890
Vaginal Intercourse without a Condom	31.8	41.2	-9.4	.009*
Anal intercourse without a Condom	5.6	7.0	-1.4	.445
Oral Intercourse without a Condom	41.7	48.5	-6.8	.066
Pregnancy	18.4	19.7	-1.3	.653
Sample size	377	342		

Note: The sample size reported here is for youth who provided a baseline and 12-month follow-up survey. Fewer cohorts of youth were eligible for the 12-month follow-up. Not all youth provided data on the four outcomes of interest and subsequently are based off of the following smaller sample sizes. Vaginal intercourse without a condom sample size was 373 treatment and 335 comparison. Anal intercourse without a condom sample size was 357 treatment and 331 comparison. Oral intercourse without a condom sample size was 357 treatment and 327 comparison. Pregnancy sample size was 351 treatment and 318 comparison.

Appendix H: Impact Model Specification

All data were analyzed using SPSS statistical software package. Program impacts between treatment and comparison were determined based on the following linear probability model:

$$(Y=1)=\beta_0+\beta_1*T+\sum\beta_k*X_k+\sum\beta_p*X_p$$

where $(Y=1)$ represents that the outcome of interest (i.e., vaginal intercourse without a condom, anal intercourse without a condom, oral intercourse without a condom and pregnancy) was observed at the focal follow-up data point being analyzed, β_1 is the difference in the observed prevalence rates of the outcome (in percentage points) across the treatment and control conditions, *T* is a treatment indicator variable (i.e., assignment to the *Crossroads* program intervention group), X_k is a vector of dummy variables representing the randomized cohorts (total number of cohorts – 1), and X_p is a vector of p covariates. X_k is the vector that contains the cohort dummy variable, which are included to gain precision from the stratified random assignment design. In this model, the impact estimate β_1 represents the impact of the intervention at the focal follow-up time period (i.e., 3 months, 6 months, and 12 months).

This model includes the baseline measurement of the outcome of interest as a covariate, to obtain a more precise impact estimate of the treatment effect. We include the other baseline covariates (i.e., age, race/ethnicity, gender, grade, and sexually active) to improve precision in the impact analysis. Only findings with $p\langle 0.05$, two-tailed test will be considered statistically significant.

Appendix I: Sensitivity analyses

In order to further assess the robustness of the impact effects additional analyses were conducted using the benchmark approach with the removal of covariates (i.e., gender, age, race/ethnicity, grade, ever been sexually active at baseline, and baseline outcome of interest), and logistic regressions with and without covariates. Cohort dummy variables were included to gain precision from the stratified random assignment design in all models. The results aligned with the findings of the benchmark analyses across all outcomes.

With the removal of covariates for the primary outcome, engaging in vaginal intercourse without a condom at three month follow-up, the results were not significant (p = .385). This finding is in alignment with the benchmark analysis for this time point. Similarly, the results of both logistic regression models were not significant (p = .932 with covariates and p = .470 without covariates). There were no significant findings for either of the secondary research questions assessed at the 3-month follow-up using the benchmark approach without covariates or with either logistic regression model, which is consistent with the findings of the benchmark analysis. A summary of these findings is provided in Table I.1 and Table I.2.

Table I.1. Sensitivity of impact analysis using data from 3 month follow-up to address the primary research question

	Benchmark Approach: Linear probability model with covariates		Sensitivity Approach 1: Linear probability model without covariates		Sensitivity Approach 2: Logistic regression model with covariates		Sensitivity Approach 3: Logistic regression model without covariates	
Treatment compared to Comparison	Diff (SE)	p-value	Diff (SE)	p-value	Odds ratio	p-value	Odds ratio	*p*-value
Vaginal intercourse without condom	-1.7 (.030)	.552	-0.03 (.034)	.385	0.987	0.932	0.907	0.470

Table I.2. Sensitivity of impact analyses using data from 3 month follow-up to address the secondary research questions

	Benchmark Approach: Linear probability model with covariates		Sensitivity Approach 1: Linear probability model without covariates		Sensitivity Approach 2: Logistic regression model with covariates		Sensitivity Approach 3: Logistic regression model without covariates	
Treatment compared to Comparison	Diff (SE)	p-value	Diff (SE)	p-value	Odds ratio	p-value	Odds ratio	*p*-value
Anal intercourse without condom	-1.0 (.017)	.573	-1.1 (.017)	.527	1.004	.883	0.961	.881
Oral intercourse without condom	1.7 (.030)	.581	.80 (.035)	.817	1.177	.295	1.111	.428

For the secondary outcomes assessed at the 6-month follow-up, vaginal intercourse without a condom remained significant with the removal of covariates $(p\langle.05)$. The results from both logistic regression models were also significant $(p\langle.05)$, which further support the findings in the benchmark analysis that indicate membership in the treatment group at the 6-month follow-up may have contributed to a decrease in risky sexual behavior. Anal intercourse without a condom and oral intercourse without a condom did not produce significant results with the removal of covariates from the linear probability model at the 6-month follow-up (p = .881 and p = .488, respectively). Similar results were found for both outcomes using the logistic regression model approach with and without covariates. A summary of these findings is provided in Table I.3.

Table I.3 Sensitivity of impact analyses using data from 6 month follow-up to address the secondary research questions

	Benchmark Approach: Linear probability model with covariates		Sensitivity Approach 1: Linear probability model without covariates		Sensitivity Approach 2: Logistic regression model with covariates		Sensitivity Approach 3: Logistic regression model without covariates	
Treatment compared to Comparison	Diff (SE)	p-value	Diff (SE)	p-value	Odds ratio	p-value	Odds ratio	*p*-value
Vaginal intercourse without condom	-8.7 (.031)	.006	-10.0 (.034)	.004	0.7	.022	0.677	.006
Anal intercourse without condom	0.3 (.018)	.866	0.3 (.019)	.881	1.065	.820	1.019	.942
Oral intercourse without condom	2.3 (.032)	.477	2.5 (.036)	.488	1.126	.444	1.074	.604

For the secondary outcomes assessed at the 12-month follow-up, no significant findings were found across outcomes, which is in alignment with the benchmark analyses. A summary of these findings is provided in Table I.4.

Table I.4. Sensitivity of impact analyses using data from 12 month follow-up to address the secondary research questions

	Benchmark Approach: Linear probability model with covariates		Sensitivity Approach 1: Linear probability model without covariates		Sensitivity Approach 2: Logistic regression model with covariates		Sensitivity Approach 3: Logistic regression model without covariates	
Treatment compared to Comparison	Diff (SE)	p-value	Diff (SE)	p-value	Odds ratio	p-value	Odds ratio	*p*-value
Vaginal intercourse without condom	-2.7 (.034)	.419	-5.7 (.037)	.124	0.942	0.724	0.794	0.131
Anal intercourse without condom	-0.6 (.021)	.776	-0.9 (.022)	.659	0.936	0.818	0.857	0.563
Oral intercourse without condom	-0.1 (.036)	.972	-2.3 (.038)	.543	1.068	0.691	0.942	0.700
Pregnancy	2.3 (.026)	.378	1.8 (.032)	.559	1.268	0.315	.1.109	0.572