Evaluation of
It's Your Game…Keep It Real
in South Carolina

Final Impact Report for

South Carolina Campaign to Prevent Teen Pregnancy

December 7, 2015

Prepared by

ETR, External Evaluator

Karin K. Coyle, PhD
Susan C. Potter, MS
Jill R. Glassman, PhD, MSW
Liz McDade-Montez, PhD
Tracy Unti, BS

Table of Contents

EVALUATION OF IT'S YOUR GAME…KEEP IT REAL IN SOUTH CAROLINA: FINDINGS FROM THE REPLICATION OF AN EVIDENCE-BASED TEEN PREGNANCY PREVENTION PROGRAM

I. Introduction

A. Introduction and study overview

Despite decreases over the last decade, teen birth rates in South Carolina (SC) are consistently higher than national rates. In 2013 South Carolina had the 12th highest teen birth rate in the country among 15 to 19 year olds, with a birthrate of 31.6 per 1,000 compared to 26.6 per 1,000 nationally (CDC, 2014). Moreover, 2013 teen birth data indicate that nearly all SC counties (85%) had teen birth rates higher than the national rate of 26.6 per 1,000 for 15-19 year olds (SC DHEC, 2014). Even greater risks exist among African American females in SC; while comprising 34% of the teen female population, African Americans in this age cohort accounted for 42% of teen births (SC DHEC, 2014).

Increasing rates of sexual activity combined with declining contraceptive use rates create a dangerous situation for youth in South Carolina. In the state, 19% of middle school students have had sex, a number that increases to 31% by the end of 9th grade and to 61% by the end of 12th grade (SC DOE, 2014). In 2013, 31% of sexually active SC middle school students did not use a condom at last intercourse (SC DOE, 2014).

These trends clearly support the importance of providing pregnancy prevention programs in middle school, before most students become sexually active. While implementation of evidence-based programs in SC middle schools is currently very limited, the state's Comprehensive Health Education Act (CHEA) mandates that certain topics be covered in middle school health education including reproductive health and the prevention of sexually transmitted infections (STIs). When this study first began, the SC Campaign to Prevent Teen Pregnancy (SC Campaign) reviewed the existing evidence-based pregnancy prevention programs for middle school (Goesling, Colman, Trenholm Terzian, & Moore, 2013). One program, called It's Your Game…Keep it Real (IYG), addresses all eight of South Carolina's health education standards and meets CHEA standards regarding mandated comprehensive reproductive health education in public schools. Consequently, the SC Campaign selected IYG for replication in selected middle schools throughout the state.

IYG has been tested in two separate studies (Tortolero, Markham, Peskin, et al., 2010; Markham, Tortolero, Peskin, et al., 2012). The first study (Tortolero et al., 2010) used a cluster randomized controlled trial design with 10 Texas (TX) urban middle schools with low-income populations; half received the 2-year intervention (12 lessons in 7th and 12 lessons in 8th). Investigators defined and tracked a cohort of 981 7th grade youth through the end of 9th grade, with 92% completing the 9th grade follow-up survey. The primary outcome variable, sexual initiation was defined as initiation of vaginal, oral or anal intercourse. Results showed that students in the comparison schools were 1.29 times more likely to initiate vaginal, oral, or anal sex by 9th grade than those in the intervention schools, and this difference reached statistical significance (p<.05). Results focusing on initiation by type of sexual intercourse showed that the intervention had a statistically significant impact on delaying oral sex (p<.01) and anal sex (p<.01); the effects for vaginal sex did not reach statistical significance for the total sample, but did for Latino students only (p<.05) . The program also reduced the frequency of vaginal

intercourse in the past 3 months (p<.05). This study met the review criteria for a moderate quality rating (Goesling, Colman, Trenholm, Terzian, & Moore, 2014).

The second study (Markham et al., 2012) used a cluster randomized controlled trial design with 15 urban middle schools; schools were assigned to one of three intervention conditions: IYG (referenced as a risk reduction program in the article), a risk avoidance program, or control. A cohort of 1,742 7th grade students was tracked into 9th grade, with 76.5% completing the 9th grade follow-up survey; the final analysis sample included 1,258 youth. The primary outcome variable, sexual initiation, was defined as initiation of vaginal, oral or anal intercourse, consistent with the first study. Results showed that students in the risk reduction condition (IYG) were less likely to initiate any type of sex (p<.01) or vaginal sex (p<.05) relative to students in the comparison schools; students receiving IYG were also less likely to report unprotected sex at last intercourse (p<.05), and reported lower frequency of vaginal (p<.05) and anal (p<.01) sex in the past 3 months, and unprotected vaginal sex in the last 3 months (p<.05). This study met the review criteria for a moderate quality rating (Goesling, Lee, Lugo-Gil, & Novak, 2014).

This report describes the implementation and impact of a replication of IYG in SC middle schools funded through a grant from the Office of Adolescent Health to the South Carolina Campaign to Prevent Teen Pregnancy. ETR was contracted to conduct the evaluation.

B. Primary research question(s)

The primary research question addressed overall program impact on sexual initiation: *What is the impact of the IYG program relative to the standard of care on initiation of vaginal sex by the end of 9th grade (12-18 months post-program) on students reporting "No" to ever had vaginal sex at baseline?* Although the original studies of IYG assessed the impact on a combined outcome of vaginal, oral or anal sex, survey questions addressing oral and anal sexual behaviors were not acceptable to SC middle schools participating in the study, and thus, this study focuses on vaginal sex.

C. Secondary research question(s)

Secondary behavioral outcomes assessed impacts on sexual initiation at an earlier point in time and addressed protected vaginal sex and number of partners: (1) *What is the impact of the IYG program relative to the standard of care on initiation of vaginal sex by the end of 8th grade (0-6 months post-program) on students reporting "No" to ever had vaginal sex at baseline?* (2) *What is the impact of the IYG program relative to the standard of care on whether (within the past 3 months) students report having had sex as measured at the 9th grade follow-up?* And *(3) What is the impact of the IYG program relative to the standard of care on whether (within the past 3 months) students report having had sex without effective birth control as measured at the 9th grade follow-up?*

II. Program and comparison programming

A. Description of program as intended

It's Your Game…Keep It Real is a two-year intervention that consists of 24 50-minute lessons, 12 delivered in 7th grade and 12 delivered in 8th grade. It was developed using a systematic instructional design process, Intervention Mapping (IM), to ground its content in

social cognitive theory, social influence models, and the theory of triadic influence (Tortolero et al., 2010). IM describes the process of health promotion program development in six steps, following the Intervention Map, and using the core processes: (1) the needs assessment, (2) the definition of proximal program objectives based on scientific analyses of health problems and problem causing factors, (3) the selection of theory-based intervention methods and practical strategies to change determinants of health-related behavior, (4) the production of the program components and production, (5) the anticipation of program adoption, implementation and sustainability, and (6) the anticipation of process and effect evaluation.

In each grade, the program integrates group-based classroom activities with personalized journaling and individual, tailored, computer-based activities. A life skills decision-making paradigm (Select, Detect, Protect) underlies the activities, teaching students to select personal limits regarding risk behaviors, to detect signs or situations that might challenge these limits, and to use refusal skills and other tactics to protect these limits. Students are taught to avoid a risky situation by either using a clear "No" or alternative action (e.g., "My parent is calling me, I have to go."). These avoidance strategies are reiterated in the curriculum activities (such as role plays and journaling activities) and computer activities. The curriculum also includes three parent-child homework activities at each grade level designed to facilitate dialogue on topics including friendship qualities, dating, and sexual behavior.

In this study, IYG lessons were intended to be delivered in a variety of classroom instructional settings (e.g., physical education, health course, or social studies). Facilitators had to be employed by the district and were required to complete a two-day training for each grade level (7th and 8th) conducted by the curriculum developers. If a facilitator was unable to attend, a trained technical assistance (TA) specialist provided comparable one-on-one training. The lessons were to be delivered during regular classroom time according to the schedule that worked at each participating school (e.g., twice a week, once a week, or every day) with no more than two weeks between lessons. Schools were allowed to teach participating students throughout the school year. For example, some schools taught half of the students in the fall semester and the other half in spring. Group size for IYG lessons was allowed to vary depending on the number of students enrolled in the classroom. During the evaluation study in SC, IYG served as the primary source for reproductive health content in the 12 intervention schools, substituting for any prior reproductive health education.

B. Description of counterfactual condition

No systematic alternative program was offered in the 12 comparison middle schools during the period the intervention schools were teaching the 7th and 8th graders enrolled in the study, including any evidence-based or promising programs. In fact, schools were not considered eligible for participation in the study if an evidence-based Teen Pregnancy Prevention (TPP) program or a promising program was being implemented or there were plans to do so during the study time frame. These criteria minimized the chance that the evaluation design would be compromised by competing programs.

The CHEA requires public schools in South Carolina serving grades 6 through 8 to cover certain health education topics, including reproductive health and STI prevention. At its core, the state's CHEA, which guides all sexuality education instruction, emphasizes local control of content; thus, school districts have the authority to implement it with varying levels of fidelity.

Each school in the comparison condition provided its usual health and sex education program, which varied by district, and included activities that addressed some or all of the following topics: puberty-reproductive health, healthy relationships, decision making (general health), decision making (sexual health), communicating values about sex, identifying and avoiding risky situations, teen pregnancy, HIV/AIDS/STIs, abstinence, condoms and contraception, media influences, and dating violence.

III. Study design

A. Sample recruitment

1. School Sample

This study involved working with selected school districts and schools throughout South Carolina. The SC Campaign recruited schools via school district administrators during the year prior to commencing the evaluation. Invitation letters were sent to all school districts; 15 school districts representing 45 schools were screened for participation. Participating districts had to meet the following criteria:

- be a public school district;
- include schools with 7th and 8th grades;
- be willing to participate and agree to the conditions of the study;
- provide the study team with school-level statistics needed for the randomization process, if not available on the SC Department of Education (SC DOE) website; and
- approve the IYG curriculum through the mandated process, (i.e., obtaining approval from the Comprehensive Health Education Committee, school improvement council, and the school board).

Participating schools had to meet the following criteria:

- be a mainstream school (not an alternative or special education school);
- include 7th and 8th grades;
- have at least 20 7th graders;
- be willing to participate and agree to the conditions of the study;
- not be involved in another federally funded project with the SC Campaign;
- not currently be using an evidence-based teen pregnancy prevention program in 7th or 8th grades; and
- not intend to begin implementing an evidence-based teen pregnancy prevention program in 7th or 8th grades in the next three years

Of the 15 school districts that were screened, 13 school districts representing 30 eligible schools agreed to participate in the study. Of the 13 school districts, the SC Campaign accepted memoranda of understanding (MOUs) from the first 24 schools representing 11 school districts at that time, which reduced to 10 after the study started because 2 of the districts merged 1 year after randomization.

All 24 schools are rural middle schools across South Carolina with total enrollments ranging from 213 to 1,486 students at the time of randomization (SC DOE, 2011). In 2011, 23.6% of South Carolinian school-aged children (5-17 years) lived in poverty, whereas the average percent

across participating school districts was 31.5%; the range was 22.4% to 42.7%. Further, at the time of randomization, the percent of students who qualified for free lunch across the 24 participating schools was 62.6%, ranging from 34.2% to 90.5% (US Census, 2011). The state of South Carolina, although racially diverse, is predominately white (67%). At the time of randomization, 11 of the 24 participating schools had more than 50% white students and 10 of the 24 schools had more than 50% black students (with 7 of those 10 having more than 70% black students). The remaining schools were mixed between white, black, Latino, and other races/ethnicities (SC DOE, 2011.)

2. Youth Sample

Youth were eligible to participate in the study if they were enrolled in 7th grade at a participating school in fall 2011, did not have limited capabilities or special needs as determined by the school, and spoke English well enough to understand the survey questions if they were read aloud. Active parental consent (i.e., the permission form must be signed by a parent/legal guardian and returned to the school) was obtained prior to data collection and at one time for all study activities. Consent forms were sent to parents via the schools' regular communication channels (e.g., weekly envelopes or parent folders). Teachers who had 90% or greater return of completed consent forms (regardless of whether consent was given or denied) received a $50 gift certificate acknowledging their support. Additionally, students who returned a consent form received a $5 gift card for returning the form, regardless of whether their parents said "Yes" or "No" to survey participation.

Because parent consent was obtained after randomization, several steps were taken to keep parents "blind" to their children's school's intervention condition. Only select district and school administration staff, school project site coordinators, and IYG facilitators at intervention schools knew their schools' condition before obtaining active parental consent and administering the baseline survey. (IYG facilitators knew because they were trained to implement IYG in August 2011.) District and school staff were explicitly asked to not share intervention condition information with anyone. To further minimize the likelihood that students or parents would learn their condition, exactly the same evaluation parental consent form was used at all 24 schools, and schools were instructed to keep the distribution processes separate from program consent for their reproductive health programming. In addition to consenting to the evaluation, parents received consent forms for programming. The programming consent forms differed slightly between the intervention and comparison conditions but were similar to diminish the likelihood that parents or students at intervention schools would associate the IYG program with the study. In intervention schools, explicit directions were provided to not tell students and teachers that IYG was being evaluated. To the evaluator's knowledge, no parent or other school staff asked or learned about their school's intervention condition during the consent process.

Parent consent return was promoted throughout the baseline data collection period. For example, if a school had low parent consent return rates the day of the scheduled baseline survey, ETR staff would make announcements to students not taking the survey, reminding them that they could still return the form and complete the survey during survey make-up sessions. By October 7, 2011, 79.9% of the 4,276 eligible students returned parent consent forms; 68.5% had positive parent consent. Student assent was obtained from all students with parental consent immediately prior to administering the survey, and it covered the entire study period. Sixty-six percent of eligible students (n=2,826) completed the baseline survey in fall 2011.

A second round of consent/assent form collection and surveying was scheduled for January-February 2012 in selected schools (6 comparison schools) due to low parent consent return rates. The procedures were similar to those used in fall 2011 with two key exceptions: (1) iPod raffles were added to the incentive structure and (2) verbal consent procedures were initiated for students who had not returned consent forms after multiple reminders over a designated period of time. As part of the verbal consent process, trained staff used a scripted protocol to obtain verbal consent from students' parents or legal guardians via telephone. The additional effort yielded an overall parent consent return rate of 89.7% (75.9% were positive) and a 73.5% student participation rate (n=3,143). In total, 76.0% of eligible youth in intervention schools (n=1,725) and 70.7% of eligible youth in comparison schools (n=1,418) completed a baseline survey between late August 2011 and early February 2012.

B. Study design

The study involved an experimental group-randomized trial design in which the 24 participating schools were randomized to receive IYG (intervention condition) or serve as comparison sites. Students were surveyed at baseline prior to the program (the fall semester of the 7th grade), post-program in the spring semester of 8th grade, 0-6 months after the program ended, and a final time in the spring semester of 9th grade which was approximately 12-18 months post-program. The amount of time between the end of the program and each follow-up survey varied across students because IYG was implemented during the semester (spring or fall) and at a frequency (e.g., 2x/week) that worked best for each school.

Given the relatively small number of units (i.e., schools) to be randomized, ETR used a restricted randomization procedure to optimize the balance of key demographic/academic characteristics shown in the literature to be related to the primary sexual initiation outcome. In particular, the procedure was an adaptation of the multi-attribute utility measurement approach (MAUM; Graham et al., 1984) and has some similarities to minimization randomization procedures (e.g., Scott et al., 2002; Pocock and Simon, 1975). In the randomization procedure, the differences between intervention and comparison schools were minimized on selected key school-level characteristics available from the SC DOE website including school configuration, racial/ethnic distribution, free lunch eligibility, and percent of students who passed standardized tests in the year prior to random assignment. Randomization was conducted prior to collection of baseline data. More detail regarding the restricted randomization procedures is provided in detail in Appendix A.

C. Data collection

1. Impact evaluation

The primary source of data for the outcome analyses was a student self-report survey. Students in both the intervention and comparison conditions were surveyed 3 times as noted above on their knowledge, attitudes, skills, intentions and behaviors related to adolescent sexuality and pregnancy. At each time point, students in intervention and comparison schools were surveyed during the same time frame, with the exception of the sub-sample of students from 6 comparison schools at baseline, as described under Sample Recruitment. See Appendix A for more details.

Data were collected by trained data collectors in school using audio-enhanced computer assisted surveys via tablets. At baseline, incentives were provided for the return of a completed parent consent form only. No monetary incentives were provided for completing the survey, but students were allowed to keep the headsets they received for completing the survey. At follow-up, students surveyed in a school setting received a $5 gift card and could keep the headsets. Students who were no longer enrolled in their original study schools were tracked and surveyed in one of several ways: (1) at their current school (first priority), (2) using an online survey or by-mail survey (second priority), or (3) using an abbreviated telephone survey (third priority). Students who completed a follow-up survey online, by mail, or by phone received $10 gift cards as an acknowledgment of using their personal time to participate in the study.

2. Implementation evaluation

Implementation data were collected from a number of different sources at different times throughout programming. Implementation logs, created by IYG developers to measure fidelity and translated to an online format by ETR, were completed by IYG facilitators on an ongoing basis. Gift card incentives were used to encourage the submission of logs within 2 school days of teaching an IYG lesson. Observations conducted by trained evaluation staff assessed both fidelity and quality of implementation; 5% of implemented lessons were observed. To obtain as representative a sample as possible data collectors observed each IYG facilitator (1) at least 2 times, and (2) a total number of times proportionate to the number of classes s/he taught. Additionally, to represent all lessons while focusing on those involving greater teacher facilitation, data collectors attempted to observe (2) each computer lesson at least 2 times, (3) each role play lesson at least 6 times, and (4) all other "regular" lessons at least 4 times. Ultimately, however, the observation sample was one of convenience due to facilitators' availability and thus, this measure may not be representative of all possible interactions. Dosage data (i.e., program attendance) were submitted by facilitators at the end of the 12 lessons for each class of students. Implementation log and observation data were reviewed by IYG project staff on a weekly basis allowing for them to provide "real time" TA to facilitators as needed (Kershner, Flynn, Prince, et. al, 2014). Project staff then recorded all TA contacts in an electronic database making it easy to identify any significant adaptations, if and when they occurred. IYG facilitators completed online reaction surveys at the end of each school year in which they provided information about their training and background. Health educators at comparison schools and those teaching the study cohort in its 9th grade year also completed online surveys asking about the content of and time spent implementing any sexual health education to the study cohort. IYG facilitators, comparison school health educators, and 9th grade health educators received $10 gift cards for completing the end-of-year surveys. See Appendix B for more details.

D. Outcomes for impact analyses

The indicators used to measure the primary behavioral outcomes are described in Table III.1 and indicators for secondary behavioral outcomes are described in Table III.2.

Table III.1. Behavioral outcomes used for primary impact analyses research questions

Outcome name	Description of outcome	Timing of measure relative to program
Initiation of sex[a]	The variable is a yes/no measure of whether the student initiated vaginal sex between baseline and the 9th grade follow-up survey. It uses the following survey item: *Have you ever had sexual intercourse?*	Spring of 9th grade (second follow-up)[b]

[a] All behavioral measures refer to vaginal sex only. The survey used the term *sexual intercourse,* which was defined as *a male putting his penis into a female's vagina*.

[b] Because intervention schools could implement IYG according to what worked best for their schedules/calendars and in either fall or spring semester, the second follow-up (spring of 9th grade) occurred 12 to 18 months post-program.

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

Outcome name	Description of outcome	Timing of measure relative to program
Initiation of sex[a]	The variable is a yes/no measure of whether the student initiated vaginal sex between baseline and the 8th grade follow-up survey. It uses the following survey item: *Have you ever had sexual intercourse*	Spring of 8th grade (first follow-up)[b]
Had sex in the past 3 months[a]	This variable is a yes/no measure of whether the student had vaginal sex in the past 3 months: *In the past 3 months, have you had sexual intercourse, even once?* In order to include the full sample of respondents in the analysis, virgins were given the value of 0.	Spring of 9th grade (second follow-up)[b]
Had sex[a] without effective birth control in the past 3 months	The variable is a yes/no measure of whether the student had vaginal sex without using effective birth control in the past 3 months. It uses the following survey item: *In the past 3 months, have you had sexual intercourse <u>without</u> you or your partner using any of these methods of birth control? [(Condoms, Birth control pills, The shot (Depo Provera), The patch, The ring (NuvaRing), IUD (Mirena or Paragard), Implant (Implanon)]*. In order to include the full sample of respondents in the analysis, virgins and those who reported not having sex in the past 3 months were given the value of 0.	Spring of 9th grade (second follow-up)[b]

[a] All behavioral measures refer to vaginal sex only. The survey used the term *sexual intercourse,* which was defined as *a male putting his penis into a female's vagina*.

[b] Because intervention schools could implement IYG according to what worked best for their schedules/calendars and in either fall or spring semester, the first follow-up (spring of 8th grade) occurred 0-6 months after the end of IYG programming and the second follow-up (spring of 9th grade) occurred 12 to 18 months post-program.

E. Study sample

Twenty-four schools were recruited into the study; all 24 schools remained in the study for its duration. The schools were randomly assigned to condition prior to the baseline survey due to logistical considerations for IYG facilitator training and to ensure sufficient time for IYG implementation after the baseline survey. The denominator for sub-cluster attrition is the number of eligible students enrolled at participating schools at the time of assignment – 2,269 students at intervention schools and 2,007 at comparison schools. The final baseline sample consisted of 1,725 students at intervention schools and 1,418 students at comparison schools; 76.0% and 70.7%, respectively, for an overall participation rate of 73.5%. Of the eligible students, 65.9% of intervention and 63.0% of comparison students completed the first follow-up survey post-program in the spring of their 8th grade year (86.7% and 89.1% of those taking a baseline survey, respectively). At the second follow-up in the spring of 9th grade, 59.8% of intervention

and 56.3% of comparison students completed a survey, representing 78.7% and 79.7% of those taking a baseline survey. By definition, the primary sample was comprised only of those who reported being virgins at baseline – 1,579 intervention students and 1,264 comparison students. Of these, 1,241 intervention and 1,027 comparison students completed a 9th grade follow-up survey and provided responses for covariates included in analyses, representing 78.6% and 81.3%, respectively, of baseline virgins and 56.9% and 58.7%, respectively, of the eligible population. Similarly for the 8th grade follow-up, 1,370 intervention and 1,131 comparison students completed a survey and provided responses for covariates included in analyses, representing 86.8% and 89.5% of baseline virgins, respectively, and 60.4% and 56.4%, respectively, of the eligible population. To improve the accuracy of estimates for the secondary behavioral outcomes, students who provided illogical responses over time to the "Ever had sex" question, i.e., responded "Yes" at baseline but "No" at follow-up, were excluded from analyses. This lowered the analytic sample for the secondary behavioral outcomes to 1,333 students in the intervention condition (58.7%) and 1,101 in comparison (54.9%). Finally, students missing values for covariates included in analysis models were excluded, yielding a final secondary analytic sample comprised of 1,307 intervention (76% of the baseline sample and 57.6% of the eligible sample) and 1,085 comparison (77% of the baseline sample and 54.1% of the eligible sample) students. See Appendix C for more detail.

F.   Baseline equivalence

The following selected variables were assessed for equivalence of the intervention and comparison conditions at baseline because literature indicates they are related to risky sexual behavior (Kirby and Lepore, 2007): age, gender, race, language spoken in the home, two indicators of home structure, academic grades, educational aspirations, two indicators of religiosity, frequency of alcohol use in the last 30 days. For the secondary sample equivalence was also assessed on the baseline values of ever had sex, had sex in the past 3 months, and had sex without using effective birth control in the past 3 months. Multilevel regression analyses were conducted with the variable of interest as the dependent variable, the intervention indicator as the independent variable, and clustering standard errors at the school level. The conditions were considered not equivalent on a given variable if the p-value was less than or equal to .05 using the Wald test. Tables III.3, III.4, and III.5 show the results of these baseline equivalence analyses for the primary and secondary analytic samples.

Table III.3. Summary statistics of key SC IYG Student Survey baseline measures for primary sample (initiation of vaginal sex, 9th grade follow-up)[a]

Baseline measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison p-value of difference[b]
Age (in years)	12.6 (0.4)	12.8 (0.5)	0.2	0.001
Gender (% female)	55.5	54.8	0.7	0.863
Race/ethnicity: Black	35.1	37.7	2.6	0.708
Race/ethnicity: Hispanic/Latino(a)	11.2	11.4	0.2	0.765
Race/ethnicity: White	42.5	41.8	0.7	0.867
Race/ethnicity: Other	11.2	9.2	2.0	0.117
English is main language spoken in the home (% yes)	88.3	89.0	0.7	0.676
Live in another home some of the time (% yes)	26.2	26.5	0.3	0.963
Mother was a teen parent (% yes)	25.1	30.8	5.7	0.219
Number of biological parents in the home (0-2)	1.42 (0.61)	1.36 (0.63)	0.06	0.281
Academic grades (1=Mostly D's and F's to 4=Mostly A's and B's)	3.54 (0.66)	3.40 (0.74)	0.14	0.101
Educational aspirations (1=Does not plan to finish high school to 6=Plan to complete a graduate degree)	5.11 (1.31)	5.06 (1.32)	0.05	0.773
Importance of faith (1=Not at all important to 4=Very important)	3.05 (0.87)	3.11 (0.87)	0.06	0.431
Number of times went to a religious service in past 12 months (1=Never to 6=Once a week)	4.36 (1.54)	4.36 (1.55)	0.0	0.908
Number of days in last 30 had 1 or more drinks of alcohol (1=0 days to 7=All 30 days)	1.17 (0.64)	1.21 (0.69)	0.04	0.206
Sample size	1,241	1,027		

[a] The primary analytic sample was comprised of students who completed a baseline survey, a 9th grade follow-up survey, provided values for covariates included in the final analysis models, and reported being a virgin at baseline.

[b] The p-values are adjusted for clustering at the level of random assignment.

Table III.4. Summary statistics of key SC IYG Student Survey baseline measures for secondary sample (initiation of vaginal sex, 8th grade follow-up)[a]

Baseline measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison p-value of difference[b]
Age (in years)	12.6 (0.4)	12.8 (0.5)	.2	0.000
Gender (% female)	56.2	54.1	2.1	0.554
Race/ethnicity: Black	35.5	37.9	2.4	0.694
Race/ethnicity: Hispanic/Latino(a)	11.2	12.2	1.0	0.998
Race/ethnicity: White	41.4	40.0	1.4	0.887
Race/ethnicity: Other	11.9	9.5	2.4	0.070
English is main language spoken in the home (% yes)	88.1	88.7	0.6	0.764
Live in another home some of the time (% yes)	26.4	27.2	0.8	0.892
Mother was a teen parent (% yes)	26.6	32.0	5.4	0.305
Number of biological parents in the home (0-2)	1.41 (0.61)	1.35 (0.63)	0.06	0.390
Academic grades (1=Mostly D's and F's to 4=Mostly A's and B's)	3.52 (0.66)	3.37 (0.74)	0.15	0.091
Educational aspirations (1=Does not plan to finish high school to 6=Plan to complete a graduate degree)	5.09 (1.33)	5.03 (1.34)	0.06	0.345
Importance of faith (1=Not at all important to 4=Very important)	3.06 (0.88)	3.09 (0.86)	0.03	0.443
Number of times went to a religious service in past 12 months (1=Never to 6=Once a week)	4.37 (1.53)	4.33 (1.56)	0.04	0.557
Number of days in last 30 had 1 or more drinks of alcohol (1=0 days to 7=All 30 days)	1.17 (0.63)	1.27 (0.74)	0.07	0.341
Sample size	1,370	1,131		

[a] The secondary analytic sample addressing sexual initiation by the end of 8th grade was comprised of students who completed a baseline survey, an 8th grade follow-up survey, provided values for covariates included in the final analysis models, and reported being a virgin at baseline.

[b] The p-values are adjusted for clustering at the level of random assignment.

14

Table III.5. Summary statistics of key SC IYG Student Survey baseline measures for secondary sample (had sex and had sex without effective birth control in the past 3 months, 9th grade follow-up)[a]

Baseline measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison p-value of difference[b]
Age (in years)	12.7 (0.4)	12.8 (0.5)	0.1	0.001
Gender (% female)	54.2	53.4	0.8	0.753
Race/ethnicity: Black	36.3	39.4	3.1	0.687
Race/ethnicity: Hispanic/Latino(a)	11.1	11.2	0.1	0.891
Race/ethnicity: White	41.2	40.2	1.0	0.879
Race/ethnicity: Other	11.4	9.2	2.2	0.098
English is main language spoken in the home (% yes)	88.5	89.0	0.5	0.811
Live in another home some of the time (% yes)	27.2	27.2	0.0	0.846
Mother was a teen parent (% yes)	26.0	32.3	6.3	0.203
Number of biological parents in the home (0-2)	1.41 (0.61)	1.35 (0.62)	0.06	0.326
Academic grades (1=Mostly D's and F's to 4=Mostly A's and B's)	3.52 (0.68)	3.38 (0.75)	0.14	0.096
Educational aspirations (1=Does not plan to finish high school to 6=Plan to complete a graduate degree)	5.08 (1.33)	5.04 (1.33)	0.04	0.801
Importance of faith (1=Not at all important to 4=Very important)	3.05 (0.88)	3.11 (0.85)	0.06	0.420
Number of times went to a religious service in past 12 months (1=Never to 6=Once a week)	4.36 (1.54)	4.35 (1.55)	0.01	0.972
Number of days in last 30 had 1 or more drinks of alcohol (1=0 days to 7=All 30 days)	1.19 (0.67)	1.24 (0.74)	0.05	0.230
Ever had sex (% yes)	5.2	5.6	0.4	0.834
Had sex in the past 3 months (% yes)	2.4	2.7	0.3	0.814
Had sex without effective birth control in the past 3 months (% yes)	1.0	0.8	0.2	0.659
Sample size	1,307	1,085		

[a]   The secondary analytic sample addressing recent sexual behaviors was comprised of students who completed a baseline survey, a 9th grade follow-up survey, provided values for covariates included in the final analysis

models, and whose answers to "Have you ever had sex?" were consistent across the 2 surveys, i.e., a "Yes" response was not followed by a "No."

[b]   The *p*-values are adjusted for clustering at the level of random assignment.

G.   Methods

1.   Impact evaluation

Multivariable analyses were conducted using multilevel regression analyses (also known as hierarchical or random coefficients regression) to evaluate the research questions as definitively as possible. Because the study design is composed of measurements taken from students nested within schools, it was anticipated that observations from students within the same school may be correlated to different degrees. Application of traditional regression estimation techniques, which assume independence between observations, to correlated data can lead to an underestimation of the standard error resulting in an increased probability of a Type I error, that is, a false positive (Goldstein, 1995). Therefore, multilevel regression analysis was used to model the data in the presence of this correlation, where level 1 was the student and level 2 was the school. In particular, multilevel *logistic* regression models were used for dichotomous outcomes (e.g., initiation of sex).

Each model included an indicator variable denoting intervention condition, the baseline outcome variable (when applicable), age, gender and race/ethnicity, and a set of *a priori* identified outcome-related covariates that differed at *p*<.15 between the conditions in the baseline sample of students who completed the corresponding follow-up when accounting for the clustered structure of the data. *A priori* covariates tested for possible inclusion in analysis models are all variables presented in Table III.3 with the exception of age, gender and race/ethnicity, which were always included in the model; only academic grades remained in the final models. Additionally, three of the variables used in the randomization process (school enrollment, school configuration, and potential exposure to an evidence-based program in the 9th grade) were included as covariates in the model regardless of whether they were imbalanced at baseline. The other remaining variables used in the randomization process were dropped from the models due to extremely high levels of correlations (r>.5) between these school-level indicators, or because they were already represented by individual-level demographic variables (race/ethnicity) and model parsimony was desired. Finally, an indicator representing whether the student completed their baseline survey in fall of 7th grade or in February of 7th grade was included in all models, as was a school-level covariate representing the percent of students reporting they ever had vaginal sex at baseline (excluding those students who completed baseline surveys in February). The latter was included in an attempt to control for potential environmental or normative influences that may have resulted from the unexpectedly large observed imbalance in rates of reported vaginal sex in the present study's sample of students taking a baseline survey in fall (7.4% in the intervention condition and 9.0% in the comparison condition).

Two sets of sensitivity analyses were conducted to understand (1) the influence of including the covariate representing the percent of students reporting they ever had sex at baseline, and (2) the influence of including the students who completed the baseline survey 3 months after the main sample. Exploratory layered analyses were conducted to better understand the contribution of different covariates to the findings.

All analyses were conducted using STATA 13.1, which utilizes maximum likelihood methods for fitting multilevel models.

2. Implementation evaluation

Implementation was evaluated through descriptive statistics and qualitative analysis. Results for analysis of Adherence, Quality, and Counterfactual indicators are presented as frequency counts percentages, averages, standard deviations, and/or ranges. Results for Context indicators are presented as frequency counts and qualitative data. See Appendix D for more detail.

IV. Study findings

A. Implementation study findings

1. Adherence

Sessions delivered. Across classes with complete log data (93.6% of 84 7th grade classes; 99% of the 99 8th grade classes), 100% of classrooms received all 12 sessions during 7th and 8th grade years. The average duration of each session was 52 minutes during 7th grade and 50 minutes during 8th grade, equating to 624 minutes and 600 minutes, respectively. The average frequency of sessions was every 3.4 days during 7th grade (range = every 2 to 10.6 days) and every 4.1 days during 8th grade (range = every 1.3 to 14.2 days). Individual facilitators determined the frequency with which sessions were delivered.

Content covered. Teachers delivered an average of 98% of the IYG activities within lessons across both years of the study (67 of 68 possible activities).

Dosage received. In 7th grade, students attended an average of 11.3 sessions (94% of 12 lessons), and 1% of students did not attend any sessions. In 8th grade, students attended an average of 10.2 sessions, or 84% of 12 lessons, and 11% of students did not attend any sessions.

Facilitator background and training. One hundred percent of IYG facilitators completed training in IYG and thereby met the qualifications to teach the curriculum. Seventeen facilitators implemented 7th grade IYG and 15 implemented it in the 8th grade. During 7th grade, seven schools had one facilitator teaching IYG and five schools had two facilitators delivering the curriculum. During 8th grade, nine schools had one facilitator and three schools had two delivering the curriculum. Most 7th and 8th grade facilitators were health and PE teachers (65% and 67%, respectively). Others included nurses, other teachers, support staff, and behavior intervention specialists. All facilitators (100%) received TA. A total of 805 hours of TA were provided during the two-year study period, an average of 67 hours per school.

2. Quality

During the 7th grade, 48 observations were made by two outside raters. On ratings of facilitator comfort level discussing sex-related topics, only 19 observations (39.6%) included ratings for this indicator because many of the 7th grade lessons do not explicitly talk about sex-related topics. Of those that did receive a rating, 89.5% had a score of 4 or 5 on a 5-point scale, where 1=poor, 3=average, and 5=excellent. On ratings of facilitator rapport with students, 87.5% of the 7th grade ratings were a 4 or 5. Among observations during which students asked questions (81% of observations), 90% included ratings of a 4 or 5 on facilitators' ability to address student questions.

In 8th grade, 64 observations were conducted. On ratings of facilitator comfort discussing sex-related topics, 95.3% of ratings received a 4 or 5; 96.8% of ratings were a 4 or 5 for rapport with students, and 91% of ratings were a 4 or 5 on facilitators' ability to address student questions (among the 83% of observations where questions were asked). On ratings of how actively the students participated in discussions and activities, 87.5% of 7th grade ratings and 93.8% of 8th grade ratings were a 4 or 5, where 1=little participation, 3=some participation, and 5=active participation.

3. Counterfactual Experiences

Ten of the 12 comparison schools provided sexual health education at the time comparison students in the study cohort were in 7th grade; 8 of 12 provided it when the students were in 8th grade. None of the comparison schools used an evidence-based curriculum, and only a few used a set text or curriculum. When they did, it was typically the state-level approved Decisions for Health by Holt, Rinehard, and Winston, Glencoe/McGraw-Hill's Meeks Heit Health and Wellness, or one of Glencoe-McGraw Hill's Teen Health courses for middle school. Of those comparison schools that implemented sexual health education in each grade, they reported providing an average of 440 minutes in 7th grade (in 10 of 12 schools) and 405 minutes in 8th grade (in 8 of 12 schools). (See Tables E.2 and E.3 for more details.) Students also received sexual health education during 9th grade; see the next section for details.

4. Context

During 7th grade, two facilitators from two different intervention schools indicated that their schools offered other forms of TPP programming – one in "guidance and science and family consumer science" and the second "during health classes." Similarly in 8th grade, two facilitators from two different intervention schools indicated that their schools offered other forms of TPP programming, one in "science class and family consumer science;" the other facilitator did not indicate where this other programming occurred. Although external district and school events related to teen pregnancy prevention occurred during the study period (see Table E.8 for complete list), it appears unlikely that these would affect the implementation of this study. No sites reported any substantial unplanned adaptations of the IYG curriculum.

When students in the study cohort moved onto high school, 8 high schools reported providing an evidenced-based program in 9th grade (Safer Choices), reaching approximately 52% of students coming from the comparison schools and 40% of students coming from the intervention schools. The high schools' use of Safer Choices was known at the time of randomization and equalized across condition at baseline, but one high school that served students from a feeder middle school in the intervention condition ended up not teaching Safer Choices the year the study cohort reached ninth grade, resulting in unbalanced exposure to an evidence-based program in 9th grade. The other high schools into which the students from the study cohort transferred provided an array of sexual health education (see Tables E.4 and E.5 for more details).

Additional details on all implementation evaluation findings described above can be found in Appendix E.

B. Impact study findings

1. Primary Sexual Initiation Outcome (9th grade follow-up)

In the sample of all students taking a baseline survey, the unadjusted baseline rates of *vaginal sex (ever had vaginal sex)* were 7.4% and 9.0% in the intervention and comparison conditions, respectively. In the sample of all students taking a baseline and 9th grade follow-up survey, the unadjusted baseline rates of vaginal sex were 5.4% (intervention) and 5.7% (comparison), indicating that higher risk students were more likely to drop out of the study by the 9th grade follow-up. In the sample of students with both a baseline and 9th grade follow-up survey, the unadjusted rate of vaginal sex at the end of 9th grade was 28.8% in both conditions. Finally, for the analytic sample of virgins at baseline who completed both baseline and 9th grade follow-up surveys, the unadjusted baseline rates of vaginal sex were 0% in both conditions by definition, while at 9th grade follow-up the rates were 24.7% and 24.6% in the intervention and comparison conditions respectively.

Ten covariates were entered into the final model to adjust for 1) the restricted randomization procedure and actual data collection timing, 2) potential differences in effects related to basic student-level demographic characteristics, 3) student-level variables showing baseline imbalance at $p < .15$ (i.e., academic grades, see Tables III.3, III.4, and III.5), and (4) the contextual imbalance in school-wide rates of vaginal sex based on the population of respondents who completed a baseline survey in fall 2011.

The 9th grade follow-up rates of *sexual initiation* estimated from the multilevel analysis model including all these covariates were 25.4% and 21.1% in the intervention and comparison conditions, respectively (Table IV.1); this difference was statistically significant (p=.039). Sensitivity analyses examined effects with and without key covariates (Appendix F). First, model-adjusted rates of initiation from sensitivity analyses in which the outcome model was run without the covariate representing school-wide baseline rates of vaginal sex were 25.3% and 21.7% for intervention and comparison conditions respectively (Table F.1), and the difference did not reach statistical significance (p=.171). In a second sensitivity analysis, model-adjusted rates of initiation for the sample of youth excluding those who completed baseline surveys later (in February of 7th grade) were 24.5% and 20.4% for intervention and comparison conditions, respectively; this difference was statistically significant (p=.039). Additionally, a series of layered analyses in which the covariates were entered in a stepwise manner showed that the difference between the adjusted rates of sexual initiation for young people in the intervention condition relative and to the control condition increased gradually with the addition of each covariate or set of covariates (Table F.3). For example, in the first layer, the model including the treatment indicator only showed no statistically significant difference in sexual initiation rates between conditions. As each covariate or set of covariates was entered, the difference between conditions in model-adjusted means gradually increased. Only when all covariates were entered was the difference between conditions big enough to reach statistical significance. Layered analyses conducted on the sub-sample of students who completed baseline surveys in fall 2011 showed a similar pattern to those analyses on the full sample (data not shown).

19

Table IV.1. Post-intervention estimated effects using data from SC IYG Student Survey to address the primary research questions

Outcome measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Intervention compared to comparison mean difference (*p*-value of difference)
Initiation of sex (9th grade follow-up)	25.4	21.1	4.3 (0.039)

Source: 9th grade follow-up survey, March-August 2014, administered 12-18 months after the program.

2. Secondary Sexual Initiation Outcome (8th grade follow-up)

For the secondary analytic sample of virgins at baseline who completed the baseline and 8th grade follow-up surveys, the unadjusted baseline rates of vaginal sex were 0% in both conditions by definition, and 16.2% (intervention) and 15.7% (comparison) at the end of 8th grade.

The same 10 covariates entered into the primary outcome analysis model were also entered into the secondary model. The final adjusted rates were 14.9% (intervention) and 12.9% (comparison), and did not reach statistical significance (p=.204, see Table IV.2).

Table IV.2. Post-intervention estimated effects using data from SC IYG Student Survey to address the secondary research questions

Outcome measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Intervention compared with comparison Mean difference (*p*-value of difference)
Initiation of sex (8th grade follow-up)	14.9	12.9	2.0 (0.204)
Had sex in the past 3 months (9th grade follow-up)	16.3	13.6	2.7 (0.135)
Had sex without effective birth control in the past 3 months (9th grade follow-up)	6.7	6.1	0.6 (0.586)

Sources: 8th grade follow-up survey, March-August 2013, administered 0-6 months after the program; and, 9th grade follow-up survey, March-August 2014, administered 12-18 months after the program

3. Secondary Behavioral Outcomes (9th grade follow-up)

In the secondary analytic sample of all students taking a baseline and 9th grade follow-up survey, the unadjusted rates of students reporting they *had sex in the past 3 months* were 2.4% and 2.7% at baseline and 16.4% and 16.9% at 9th grade follow-up in the intervention and comparison conditions, respectively. The unadjusted rates of students reporting they *had sex without effective birth control in the past 3 months* (*unprotected sex*) at baseline were 1.0% and 0.8% in the intervention and comparison conditions, respectively, and 7.2% and 7.6% at 9th grade follow-up.

Model-adjusted rates of students reporting *having had sex in the past 3 months* in the intervention and comparison conditions, generated from the multilevel outcome model at 9th grade follow-up, were 16.3% and 13.6%, respectively (Table IV.2). This effect was not statistically significant (p=.135). Sensitivity analyses, reported in Appendix F, show similar

results to the benchmark analyses; neither model reached statistical significant for this outcome (p=.335 and p=.115).

Finally, the model-adjusted rates of *sex without effective birth control in the past 3 months* from the multilevel multivariable outcome model at 9th grade follow-up were 6.7% and 6.1% for the intervention and comparison conditions, respectively (Table IV.2). This effect was not statistically significant (p=.586); the results of the two sensitivity analyses were similar (p=.893 and p=.499, Appendix F).

V. Conclusion

This evaluation tested a strong implementation of IYG compared to usual sexual health education in SC. Indeed, data from the implementation evaluation indicate that IYG was delivered with high fidelity to the original curriculum and with high quality across both years. Further, students attended a high percentage of lessons. Nonetheless, the behavioral results indicate that there were no statistically significant differences in rates of sexual initiation at the end of 8th grade between students who received IYG compared to those receiving the usual sexual health education programming in the comparison schools. By the end of 9th grade, study data showed that students receiving the usual sexual health education programming, which included 9th grade implementation of an evidence-based program (Safer Choices) in both conditions with differing exposure (see below), had lower rates of sexual initiation than those receiving IYG. While the difference was less than a small effect (d=.10) using Cohen's effect size standards (Cohen, 1992), it was statistically significant. The 9th grade findings also showed that there were no statistically significant differences between the two conditions on vaginal sex in the past 3 months and vaginal sex without effective birth control in the past 3 months, which included condoms.

The behavioral findings from this study differ from those in the original IYG studies (Tortolero et al., 2010 and Markham et al., 2012), which showed statistically significant reduced rates for a combined sexual initiation variable (vaginal, oral, or anal sex), as well as reduced rates of having vaginal sex and having vaginal sex without a condom in the past 3 months for the IYG condition compared to the standard of care at that time (Tortolero et al., 2010; Markham et al., 2012).

There are a number of study context and design issues that may have contributed to the pattern of findings in the SC study. First and foremost, consistent with SC health education policies, students in both conditions received strong sexual health education throughout the study period. Based on implementation data collected during the study, the topics covered in the comparison schools were similar to those covered in IYG but were not considered to be an evidence-based program because they had not yet been evaluated. Further, the average number of minutes provided for students in the comparison schools that taught sexual health education equaled approximately 7 hours per year versus 10 hours per year for IYG. In 9th grade, students in both conditions were exposed to an evidence-based high school program (Safer Choices); however, a greater proportion of the students from the comparison schools received Safer Choices in 9th grade than did students from the intervention condition. Both the nature and level of sexual health programming delivered in comparison middle schools and the disproportionate exposure to an evidence-based program in the 9th grade may account for the differential rates of sexual initiation observed at the end of 9th grade.

A number of other important factors may have contributed to the pattern of behavioral results as well. This study differed from the original IYG studies in four areas: (a) the present study was an effectiveness trial using school teachers for curriculum implementation rather than outside health educators; (b) the present study implemented the IYG program in an alternate geographic region (SC versus TX) and setting (rural versus urban) with very different racial/ethnic compositions (primary sub-groups included youth identifying as black or white in SC versus black or Latino in the original studies); (c) the present study used a different implementation schedule (schools could implement either semester during the year rather than in a specified period of time); and (d) the present study's primary outcome measure was initiation of vaginal sex only rather than the combined variable measuring initiation of oral, vaginal and anal sex used in the original IYG studies. Indeed, both original IYG studies found a positive effect on the combined outcome, but the results for the first study showed no statistically significant effects for initiation of vaginal intercourse only for the full sample (Tortolero et al., 2010).

Further, there was an imbalance across intervention and comparison schools in several individual and school-level factors known to be related to risk for sexual initiation, including school-wide rates of vaginal sex at baseline. We controlled for these to the extent possible given the data available, but the analytic models may not have fully accounted for the differences. To evaluate the primary outcome of sexual initiation, we used the subset of youth in our sample who were not yet sexually active at baseline; thus by definition the analysis dataset for this outcome had equivalent rates of sexual activity – i.e., zero – at baseline. However, the baseline data of the full sample indicated that the population of entering 7th grade students in these schools who already had initiated sex was different in the intervention and comparison conditions – 7.4% versus 9.0%, respectively. Because this difference might represent environmental or normative influences on the sexual initiation outcome – e.g., students in a school with much higher rates of sexual activity might be more likely to initiate sex sooner than students in a school with lower rates of sexual activity – the variable was included in the analytic models to control for its possible influence. To test the influence of this variable, analyses were conducted with and without this covariate. The analysis that excluded this factor did not confirm the benchmark results for the primary outcome where this factor was included.

The evaluation team collected additional baseline data 3 months after the first round of data collection to increase relatively low participation rates in 6 comparison schools, which could have affected the differential adjusted rates of sexual initiation between intervention and comparison schools. However, sensitivity analyses excluding these "late" baseline youth showed the same pattern of results suggesting this issue was not a key factor affecting the outcome results. Finally, the rates of sexual initiation at baseline in the sample of youth who took a 30-month follow-up survey were 5.4% (intervention) and 5.7% (comparison), compared to 7% and 9%, respectively, for students taking fall baseline only, indicating that more high-risk youth dropped out in the comparison condition by the end of the study; this differential attrition also could have influenced the differential adjusted initiation results seen at final follow-up.

Because this study varied on multiple factors during implementation, including setting, population, and outcome, it is difficult to tease out the effects of altering any single factor. Further replication research will continue to explore how these factors affect the robustness of the original study findings.

Future Analyses

This report represents the results of the primary and secondary behaviors only. Additional analyses are underway to examine other critical comparative intervention effects, such as on the theory-based psychosocial outcomes (e.g., knowledge, attitudes, and beliefs) as well as behavioral impacts on key sub-groups, including males versus females, and those based on race/ethnicity. The results of these analyses will be reported through a peer-reviewed journal article, as will results of mediation analyses, which can provide a better understanding of what parts of the intervention worked to influence behavior in the desired direction, and what parts did not in this population and setting.

Strengths and Limitations

This study is unique in its focus on the replication of an evidence-based program in a rural setting. The study featured a randomized design involving 24 schools and long term follow-up of youth. All schools remained in the study throughout the length of the project. As is standard practice for school-based studies, the outcome data were collected using self-report questionnaires, which are subject to potential response biases. Some evidence supports the general reliability and validity of adolescents' reports of sexual and contraceptive behaviors, particularly with the use of electronic devices (Trapl, Borawski, Stork, et al., 2005; Coyle, Russell, Shields, & Tanaka, 2007; Palen, Graham, Smith, et al., 2008). The study included youth in rural middle schools in SC; the results may not generalize to rural or urban settings in other geographic regions. We did not correct the statistical procedures for multiple testing; however, we limited the number of hypotheses and specified them *a priori*. Additionally, implementation of the study's group-randomized trial design with only 24 units of randomization (schools) resulted in some imbalances in risk and 9th grade counterfactual exposure rates across intervention and comparison conditions that affect interpretation of the results. First, randomization of schools resulted in imbalance of sexual risk across conditions at baseline, as indicated by differing rates of sexual initiation. Second, while the number of study middle schools feeding into high schools that implement an evidence-based program (Safer Choices) in 9th grade was equal across conditions at baseline, one high school that served only students from an intervention middle school ended up not teaching Safer Choices the year the study cohort reached 9th grade, which led to more comparison students receiving this program than intervention students.

This study highlights the need for further programming options and research with middle-school aged youth in rural settings. Almost none of the existing evidence based programs were developed and evaluated in rural settings, and some evidence suggests that rural youth may respond differently to programs initially developed and tested in urban settings (Goesling et al., 2013). This may be a result of a mismatch in underlying theoretical models (Elder, Guadalupe, Zabinski et al., 2001), a mismatch in content or context (e.g., failing to address the most influential risk and protective factors related to sexual behaviors in that region), or other factors, such as unaddressed cultural or regional differences.

VI. References

Centers for Disease Control (2014). *National Vital Statistics Reports: National and State Patterns of Teen Births in the United States, 1940–2013.* Retrieved January 29, 2015 from http://www.cdc.gov/nchs/data/nvsr/nvsr63/nvsr63_04.pdf

Cohen, J. (1992). A power primer. Psychological bulletin, 112(1), 155.

Coyle, K.K., Russell, L.A., Shields, J.P., & Tanaka, B.A. (2007). *Collecting Data from Children Ages 9-13*. Report prepared for the Lucile Packard Foundation for Children's Health. September, 2007.

Elder, J.P., Guadalupe, X.A., Zabinski, M.F., Prochaska, J.J., and Gehrman, C.A. (2001). Theories, Models & Methods of Health Promotion in Rural Settings. In S. Loue & B.E. Quill (Eds.). *Handbook of Rural Health* (pp. 295-314). US: Springer.

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

Goesling B, Lee J, Lugo-Gil J, & Novak T. (2014, July). Updated findings from the HHS Teen Pregnancy Prevention Evidence Review: January 2011 through April 2013. Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. Retrieved February 1, 2015 from: http://tppevidencereview.aspe.hhs.gov/pdfs/Summary_of_findings.pdf

Goldstein, H. (1995). *Multilevel Statistical Models*. London: Arnold Publishers.

Graham, J.W., Flay, B.R., Johnson, C.A., Hansen W.B., & Collins, LM. (1984). Group comparability: a multiattribute utility measurement approach to the use of random assignment with small numbers of aggregated groups. *Evaluation Review, 8*(2), 247-260.

Kershner, S., Flynn, S., Prince, M., Potter, S.C., Craft, L., & Alton, F. (2014). Using data to improve fidelity when implementing evidence-based programs. *Journal of Adolescent Health, 54*, S29-S36.

Kirby, D. (2007). *Emerging Answers 2007: New Research Findings on Programs to Reduce Teen Pregnancy — Full Report*. Washington, DC: The National Campaign to Prevent Teen and Unplanned Pregnancy.

Markham, C.M., Tortolero, S.R., Peskin, M.F., Shegog, R., Thiel, M., Baumler, E.R., Addy, R.C., Escobar-Chaves, S.L., Reininger, B., & Robin, L. (2012) Sexual Risk Avoidance and

Sexual Risk Reduction Intervention for Middle School Youth: A Randomized Controlled Trial. *Journal of Adolescent Health. (50)*3; 279-288. DOI: 10.1016/j.jadohealth.2011.07.010

Murray, D. (1998). *Design and analysis of group-randomized trials.* New York: Oxford University Press.

Murray, D.M. & Blitstein, J.L. (2003). Methods to reduce the impact of intraclass correlation in group-randomized trials. *Evaluation Review, 27*(1), 79-103.

Palen, L., Graham, J.W., Smith, E.S., Caldwell, L.L., Matthews, C., & Flisher, A.J. (2008) Rates of missing responses in personal digital assistant (PDA) versus paper assessments. *Evaluation Review, 32*(3), 257-272.

Pocock, S.J. & Simon, R. (1975). Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics, 31*(1), 103-115.

Scott, N.W., McPherson, G.C., Ramsay, C.R., & Campbell, M.K. (2002). The method of minimization for allocation to clinical trials: A review. *Controlled Clinical Trials, 23* (2002) 662–674.

South Carolina Department of Education (SC DOE), Research Portal website, 2011 Student Counts data.

South Carolina Department of Education (SC DOE), 2013 Youth Risk Behavior Survey Data.

South Carolina Department of Health and Environmental Control (SC DHEC), Division of Biostatistics and Health GIS, Office of Public Health and Statistics and Information Systems, 2013 Teen Birth Data (2014).

Tortolero, S.R., Markham, C.M., Peskin, M.F., Shegog, R., Addy, R.C., Escobar-Chaves, S.L., & Baumler, E.R. (2010) It's Your Game…Keep It Real: Delaying Sexual Behavior with an Effective Middle School Program. *Journal of Adolescent Health. 46*, 169-179. DOI: 10.1016/j.jadohealth.2009.06.008

Trapl, E.S., Borawski, E.A., Stork, P.P., Lovegreen, L.D., Colanbianchi, N., Cole, M.L., & Charvat, J.M. (2005). Use of audio-enhanced personal digital assistants for school-based data collection. *Journal of Adolescent Health, 37*, 296-305.

US Census. Small Area Income and Poverty Estimates website, 2011 school district data. Retrieved July 21, 2015 from http://www.census.gov/did/www/saipe/data/schools/index.html

Appendix A: Randomization and data collection efforts

Randomization procedure

The following seven school-level characteristics were used in the minimization process: school configuration (middle school or middle/high school), percent of students at the school who were white, percent who were black, percent who were Latino, percent who were eligible for free lunch, percent of students who passed the English language arts standardized test, and percent of students who passed the math standardized tests. First, the study statistician identified *all possible combinations of school allocations to two groups*. Second, the statistician restricted the list of possible combinations/allocations based on "stratification" variables representing district and school plans to implement a 9th-grade evidence-based curriculum addressing teen pregnancy/STI/HIV when the study cohort would be in 9th grade, eliminating combinations/allocations for which there was not balance (or close to balance in the case of odd numbers of schools per district) on these stratification variables. Third, for each of the seven characteristics and for each combination/allocation the mean differences between groups were computed. Fourth, for each characteristic the rank of this mean difference was computed, from lowest to highest. Fifth, a composite variable representing the sum of the ranks across all 7 variables was created; each variable in the composite received a weight based on the strength of the relationship the literature suggested it has with the primary sexual initiation variable. Finally, the combinations were ranked according to the sum of the ranks from lowest to highest and the combination with the smallest sum of ranks of the mean differences in key characteristics was selected as the optimal allocation of schools to two groups; one group within that allocation was randomly assigned to the intervention condition and one to the comparison condition using a table of random numbers.

Table A.1. Data collection efforts used in the impact analysis of It's Your Game…Keep It Real and timing.

Data collection effort	Cohort 1
Baseline survey [a]	8/29/2011 to 10/7/2011 1/30/2012 to 2/14/2012[b]
Start & end dates of 7th grade programming [a]	9/6/2011 to 5/21/2012
Start and end dates of 8th grade programming	8/4/2012 to 5/31/2013
First follow-up (spring of 8th grade)	3/24/2013 to 9/16/2013[c]
Second follow-up (spring of 9th grade)	3/24/2014 to 9/16/2014[d]

[a] Although there is overlap in the time frames for surveys and programming, 7th grade IYG programming did not begin at a given school until <u>after</u> the baseline survey was completed at that school. Likewise, all 8th grade IYG programming at a given school was completed <u>before</u> the 8th grade survey was administered at that school.

[b] The sub-sample of youth who completed the baseline survey in January-February 2012 were enrolled at 6 comparison schools only.

[c] School-based surveying ended on 5/31/2013; online, mail, and phone surveys continued through the summer.

[d] School-based surveying ended on 6/3/2014; online, mail, and phone surveys continued through the summer.

Appendix B: Implementation evaluation data collection

Table B.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 1-Sessions delivered: How many program sessions were offered and how often?	Web-based implementation logs (assess number of sessions, length of each session, date of sessions, among other elements) *Note*: Data on length of each session may vary by facilitator and in some instances reflect length of class period, not the minutes needed to teach. We will assess data quality before making a final decision on reporting it.	Data were collected throughout implementation on all sessions. Teachers were expected to log sessions within 2 school days of teaching them, and were incentivized to log in a timely manner.	Evaluation staff; project staff monitor data for TA needs
Adherence 2-Content covered: What content was delivered to youth?	Web-based implementation logs (specific activities completed, adaptations)	Data were collected throughout implementation on all sessions. Teachers were expected to log sessions within 2 school days of teaching them, and were incentivized to log in a timely manner.	Evaluation staff
Adherence 3-Dosage received: What and how much was received?	Attendance records *Note*: Data are limited to total number of lessons received; we do not have attendance by lesson.	Student attendance was captured in an Excel spreadsheet that was collected from IYG facilitators at the end of the 12 sessions.	Evaluation staff; project staff monitor for completion of records
Adherence 4-Faciliator background and training: Who delivered material to youth?	List of facilitators from participating schools trained to implement program, including *position in Districts* (e.g., PE teachers, nurses)	Data were collected annually, with updates throughout school year if facilitators transition.	Project staff; evaluation staff
	Record data on facilitator role *in school*, hours of TA received (by school and facilitator)	Data on TA was collected throughout year. Teacher role information was collected annually.	Project staff; evaluation staff
	Teacher survey (previous teaching experience including teen pregnancy prevention related lessons; other related training experiences)	Teacher surveys were collected annually.	Project staff; 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
Quality 1: Quality of facilitator-participant interactions	Observations of interaction quality using required OAH observation protocol (Not a direct assessment of facilitator-participant interaction, but assesses rapport and communication with participants) On the following scale (1=poor, 3=average, 5=excellent), rate the implementer on the following qualities… Comfort level discussing sex related topics e.g., reproductive anatomy, sex, condoms, contraception, teen pregnancy, STIs, etc.; Rapport and communication with participants; Effectively addressed questions/concerns	Sample of 5% of all sessions across facilitators when the evaluation cohort was in 7th and 8th grades (school years 2011-2012 and 2012-2013). *Note*: To strive for a representative sample, we attempted to observe each computer lesson at least 2 times, each lesson with role plays at least 6 times, and all other lessons at least 4 times. Additionally, each facilitator was observed at least 2 times. Ultimately, however, the sample was one of convenience due to facilitators' availability and thus, this measure may not be representative of all possible interactions.	Evaluation staff
Quality 2: Quality of youth engagement with program	Observations of engagement using OAH observation protocol How actively did the group members participate in discussions and activities? Scale: 1=little participation, 3=some participation, 5=active participation	5% of all sessions across facilitators when the evaluation cohort was in 7th and 8th grades (school years 2011-2012 and 2012-2013). *Note*: To strive for a representative sample, we attempted to observe each computer lesson at least 2 times, each lesson with role plays at least 6 times, and all other lessons at least 4 times. Additionally, each facilitator was observed at least 2 times. Ultimately, however, the sample was one of convenience due to facilitators' availability and thus, this measure may not be representative of all possible interactions.	Evaluation staff
Counterfactual 1: Experiences of comparison condition	Health teacher survey (online with a paper-pencil option for those requesting it) *Note*: Survey focuses on sexual health education (topics covered via class lessons and through other school-wide events).	Annual survey of all teachers was conducted in the counterfactual condition responsible for teaching health (7th grade completed in April 2012; 8th grade completed in April 2013).	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 1: Substantial unplanned adaptation(s)	Web-based implementation log	Log data were collected throughout implementation on all sessions. Teachers were expected to log sessions within 2 school days of teaching them, and were incentivized to log in a timely manner.	Evaluation staff; project staff monitor data for adaptations
	Record data (TA notes from meetings with sites; updated implementation plans showing substantial implementation changes)	Record data were collected ongoing through each year.	Project staff
Context 2: Other TPP programming available or offered to study participants (both intervention and comparison)	Meetings with district staff	Once per year.	Project staff
	Health teacher/IYG facilitator survey (online with a paper-pencil option for those requesting it) *Note*: Survey focuses primarily on other types of educational activities (e.g., assemblies). Could capture other TPP programming if teachers write in other TPP curricula.	Annual survey of all health teachers/IYG facilitators was conducted in both intervention and comparison conditions responsible for teaching health (7th grade completed in April 2012; 8th grade completed in April 2013; 9th grade in April 2014).	Evaluation staff; project staff
Context 3: Exposure to sexual health education in 9th grade	Teacher survey (online with a paper-pencil option for those requesting it)	One-time survey of sexual health education teachers in schools serving intervention and comparison youth when in 9th grade (April 2014)	Evaluation staff
Context 4: External events affecting implementation	Google alerts specific to study schools and Districts	Ongoing throughout the year (school and calendar year).	Project staff
	Record data (TA notes from site visits, on-site meetings, etc.)	Ongoing throughout the year (school and calendar year).	Project staff

TPP = Teen Pregnancy Prevention.

Appendix C: Study sample

Table C.1a. Cluster and youth sample sizes by intervention condition

Number of:	Time period	Total sample size	Intervention sample size	Comparison sample size	Total response rate	Intervention response rate	Comparison response rate
Clusters: At beginning of study		*24*	*12*	*12*			
Clusters: Contributed at least one youth at baseline (fall of 7th grade)	*Baseline*	*24*	*12*	*12*	*100%*	*100%*	*100%*
Clusters: Contributed at least one youth at follow-up (spring of 8th grade)	*Spring 2013 (0-6 months post-program)*	*24*	*12*	*12*	*100%*	*100%*	*100%*
Clusters: Contributed at least one youth at follow-up up (spring of 9th grade)	*Spring 2014 (12-18 months post-program)*	*24*	*12*	*12*	*100%*	*100%*	*100%*
Youth: In non-attriting clusters/sites at time of assignment		*4,276*	*2,269*	*2,007*			
Youth: Who consented (consented = had parent consent and student assent)	*Aug 2011-Feb 2012 Prior to baseline survey*	*3,143*	*1,725*	*1,418*	*73.5%*	*76.0%*	*70.7%*
Youth: Contributed a baseline survey	*Fall 2011 (7th grade)*	*3,143*	*1,725*	*1,418*	*73.5%*	*76.0%*	*70.7%*
Youth: Contributed a baseline survey and reported not having had sexual intercourse	*Fall 2011 (7th grade)*	*2,843*	*1,579*	*1,264*	*66.5%*	*69.6%*	*63.0%*
Youth: Contributed a follow-up survey (spring of 8th grade)	*Spring 2013 (0-6 months post-program)*	*2,760*	*1,496*	*1,264*	*64.5%*	*65.9%*	*63.0%*
Youth: Contributed a follow-up survey (spring of 9th grade)	*Spring 2014 (12-18 months post-program)*	*2,487*	*1,357*	*1,130*	*58.2%*	*59.8%*	*56.3%*

Number of:	Time period	Total sample size	Intervention sample size	Comparison sample size	Total response rate	Intervention response rate	Comparison response rate
Youth: Reported consistent answers across time on "Ever had sex" (i.e., a "Yes" response was not followed by a "No" at a subsequent time)	N/A	2,434	1,333	1,101	56.9%	58.7%	54.9%
Youth: Reported never having had sex at baseline and had values for all covariates that were included in final analysis model for initiation of vaginal sex at 8th grade follow-up	N/A	2501	1,370	1,131	58.5%	60.4%	56.4%
Youth: Reported never having had sex at baseline and had values for all covariates that were included in final analysis model for initiation of vaginal sex at 9th grade follow-up	N/A	2,268	1,241	1,027	53.0%	54.7%	51.1%
Youth: Had values at baseline and 9th grade follow-up for secondary outcomes assessing recent sexual activity, and had values for all covariates included in the final analysis model	N/A	2,392	1,307	1,085	55.9%	57.6%	54.1%

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 1-Sessions delivered	a. *Total number of sessions offered* is a sum of the sessions captured in the web-based project implementation log. (Note: session=lesson)
Adherence 1-Sessions delivered	b. *Average session duration* was calculated as the average of the self-reported session lengths, measured in minutes. (Note: session=lesson)
Adherence 1-Sessions delivered	c. *Average frequency of sessions* was calculated as the total number of sessions divided by the total number of days between lesson 1 and lesson 12. (Note: session=lesson)
Adherence 2-Content covered	a. *Average percentage of activities completed, per class,* was represented by the total number of activities completed with no or minor adaptations (per response categories on self-report logs) divided by the total number of activities across all 12 lessons (per the curriculum).
Adherence 3-Dosage received	a. *Average number of sessions attended* was calculated as the average of the number of sessions that each student attended, averaged across all students in the intervention condition. The maximum number possible is 12. (Note: session=lesson)
Adherence 3-Dosage received	b. *Percentage of intervention students that did not attend any IYG sessions* during the year.
Adherence 4-Facilitator background and training	a. *Count of facilitators implementing the program* across schools each year.
Adherence 4-Facilitator background and training	b. *Average number of facilitators per school* implementing the program each year.
Adherence 4-Facilitator background and training	c. *Percentage of facilitators across schools in a specified position*, using official school titles.
Adherence 4-Facilitator background and training	d. *Percentage of facilitators trained* was calculated as the number of facilitators who were trained to become certified IYG facilitators (i.e., completed a 3-day IYG training of facilitators) divided by the total number of facilitators who delivered the program.
Adherence 4-Facilitator background and training	e. *Percentage of facilitators receiving technical assistance (TA)* each year was calculated as the number of facilitators who received TA by the total number of facilitators who delivered the program each year.
Quality 1: Facilitator-participant interactions	Quality of facilitator-participant interactions was calculated as the percentage of observed where the independent evaluator scored the interaction described in the following rows as a 4 or 5 on the rating scale where 1=poor and 5=excellent.

Implementation element	Methods used to address each implementation element
Quality 1: Facilitator-participant interactions	a. Comfort level discussing sex related topics, e.g., reproductive anatomy, sex, condoms, contraception, teen pregnancy, STIs, etc.
Quality 1: Facilitator-participant interactions	b. Rapport and communication with participants
Quality 1: Facilitator-participant interactions	c. Effectively addressed questions/concerns
Quality 2: Youth engagement with program	Percent of sessions where the independent evaluator scored the following indicator as a 4 or 5 on a scale where 1=little participation and 5=active participation: a. How actively did the group members participate in discussions and activities?
Counterfactual 1	a. *Count of comparison schools providing sexual health education* each year.
Counterfactual 1	b. *Count of comparison schools using an evidence-based teen pregnancy prevention* (EBP) curriculum each year.
Counterfactual 1	c. *Count of comparison schools covering the specified sexual health topic and average number of minutes spent on that topic* each year.
Context 1: Substantial unplanned adaptation(s)	a. *Number of substantial unplanned adaptations*, including a qualitative description of any adaptations made.
Context 2: Other TPP programming available or offered to study participants (both intervention and comparison)	a. All other TPP programming available to both intervention and comparison conditions described by district personnel or health teachers/IYG facilitators via the teach survey is listed in qualitative form.
Context 3: Sexual health education in 9th grade	a. *Count of schools providing sexual health education*.
Context 3: Sexual health education in 9th grade	b. *Count of schools using an evidence-based teen pregnancy prevention curriculum*.
Context 3: Sexual health education in 9th grade	c. *Count of schools covering a specified sexual health topic and average number of minutes spent on that topic*.
Context 4: External events affecting implementation	a. *Count of schools in which implementation was affected by a specified district initiative* (unrelated to the TPP programming that occurred in this project).

TPP = Teen Pregnancy Prevention

Appendix E: Implementation study findings

Table E.1: Analysis results of implementation adherence, quality, and context at intervention schools

Implementation element	7th Grade – 2011-2012	8th Grade – 2012-2013
Adherence 1-Sessions delivered: a. total number of sessions offered	12 out of 12 sessions (100%) delivered across all classrooms. *Note.* Total n = 95.	12 out of 12 sessions (100%) delivered across all classrooms. *Note.* Total n = 99.
Adherence 1-Sessions delivered: b. average session duration	52 minutes (range = 45-62)	50 minutes (range = 33-67)
Adherence 1-Sessions delivered: c. average frequency of sessions	Every 3.4 days (range = 2-10.6)	Every 4.1 days (range = 1.3-14.2)
Adherence 2-Content covered: a. average number and average percentage of activities completed, by class, with minor or no adaptations across all 12 lessons (maximum number of activities possible is 68 in each grade)	Average number = 67 (range = 0-68) Average percent = 98.5% (range = 0%–100.0%) See Table E.6 for more details by lesson.	Average number = 67 (range = 51-68) Average = 98.5% (range = 75.0%-100.0%) See Table E.7 for more details by lesson.
Adherence 3-Dosage received: a. average number and average percentage of sessions attended (maximum possible is 12)	Average = 11.3 or 94% *Note.* 6 participants attended the full IYG curriculum twice in 7th grade. Their data is excluded in this average.	Average = 10.1 or 84% *Note.* 11 participants attended the full IYG curriculum twice in 8th grade. Their data is excluded from this average.
Adherence 3-Dosage received: b. percentage of intervention students that did not attend any IYG sessions	n = 18, 1% of total sample (n=1601) without missing data. *Note.* Data missing for 189 participants.	n=180, 11.4% of total sample (n=1576) without missing data. *Note.* Data missing for 214 participants.
Adherence 4-Facilitator background & training: a. total number of facilitators delivering program across schools each year	N = 17 facilitators	N = 15 facilitators
Adherence 4-Facilitator background & training: b. average number of facilitators per school delivering program each year	Average = 1.4, range = 1-2 per school	Average = 1.3, range = 1-2
Adherence 4-Facilitator background & training: c. facilitators positions (official school titles)	65% health & PE teachers 29% other support staff	67% health & PE teachers 27% other support staff

Implementation element	7th Grade – 2011-2012	8th Grade – 2012-2013
Adherence 4-Facilitator background & training: d. percentage of facilitators across schools trained in IYG	100%	100%
Adherence 4-Facilitator background & training: e. percentage of facilitators across schools receiving TA[a]	100%	100%
Quality of facilitator-participant interactions: Teacher's comfort with sex-related topics - Percentage of observed interactions with a score of 4 or 5 out of 5 (1=poor, 3=average, 5=excellent)	100%	100%
Quality of facilitator-participant interactions: Teacher's rapport with students - Percentage of observed interactions with a score of 4 or 5 out of 5 (1=poor, 3=average, 5=excellent)	87.5%	96.8%
Quality of facilitator-participant interactions: Teacher's ability to address questions - Percentage of observed interactions with a score of 4 or 5 out of 5 (1=poor, 3=average, 5=excellent)	72.9% Note: 18.8% responded "n/a" when there wasn't an opportunity for a facilitator to answer questions, e.g., questions weren't asked because it was a computer lesson.	75.0% Note: 17% responded "n/a" when there wasn't an opportunity for a facilitator to answer questions, e.g., questions weren't asked because it was a compute lesson
Quality of youth engagement: percentage of sessions receiving score of 4 or 5 out of 5 on group participation (1=little participation, 3=some participation, 5=active participation)	87.5% (n = 48 observations)	93.8% (n = 64 observations)
Substantial unplanned adaptation(s): number of substantial unplanned adaptations	0	0
Other TPP programming available or offered to study participants at intervention schools	2 out of the 18 facilitators surveyed reported that sex education was offered at their sites through other (non-IYG) means.	2 out of the 17 facilitators surveyed reported that sex education was offered at their sites through other (non-IYG) means.

[a] Each intervention school received an average of 67 hours of TA during the 2 years of IYG implementation, which translates into 805 hours total of TA provided by project staff from 8/1/2011 through 7/1/2013.

Table E.2. Analysis results of counterfactual at comparison schools – Overall

Implementation element	7th Grade – 2011-2012	8th Grade – 2012-2013
a. Number of comparison schools providing sexual health education	10 of 12	8 of 12
b. Number of comparison schools using an evidence based curriculum	0	0

Table E.3. Analysis results of counterfactual at comparison schools – Sexual health education topics covered

Sexual health education topic	Number of schools covering topic in 7th grade (n=10)	Average minutes[a] spent on topic in 7th grade (standard deviation, range)	Number of schools covering topic in 8th grade (n=8)	Average minutes[a] spent on topic in 8th grade (standard deviation, range)
Puberty	6	96 (79, 20-240)	5	44 (27, 20-100)
Healthy Relationships	7	46 (24, 20-80)	5	31 (12, 15-50)
Decision making for health in general	8	77 (47, 20-180)	6	36 (27, 15-100)
Decision making for sexual health	7	48 (23, 20-75)	7	39 (36, 10-100)
Communicating values about sex	6	45 (40, 10-120)	5	24 (13, 15-50)
Identifying/avoiding risky sexual situations	5	56 (43, 20-120)	7	34 (28, 15-100)
Teen pregnancy	4	33 (19, 20-60)	3	44 (18, 20-60)
HIV/AIDS and other STIs	7	129 (111, 30-360)	7	72 (89, 20-300)
Abstinence	7	77 (74, 15-120)	7	67 (89, 15-300)
Condoms and/or contraception	2	68 (74, 15-120)	3	26 (5, 20-30)
Media influence on sexual health	6	49 (43, 15-120)	5	26 (16, 10-50)
Dating Violence	4	43 (15, 30-60)	4	25 (17, 12-50)
Total minutes implementing sexual health education		439 (273, 140-960)		405 (247, 195-950)

[a] The average numbers of minutes refers only to the schools implementing that topic.

Table E.4. Analysis results 9th grade sexual health education for both intervention and comparison conditions

Implementation element	Number of high schools
a. Number of high schools (out of 20) providing sexual health education	17 (represents 11 intervention middle schools and 10 comparison middle schools feeding into these high schools)
b. Number of high schools (out of 20) using an evidence based curriculum	8 (represents 4 intervention middle schools and 5 comparison middle schools feeding into these high schools)

Note. Not all 9th graders received sexual health education even in high schools providing it.

Table E.5. Analysis results 9th grade sexual health education for both intervention and comparison conditions – sexual health topics

Sexual health education topic	Number of schools covering the topic	Average minutes[a] spent on topic (standard deviation, range)
Puberty	8	64 (77, 8-270)
Healthy Relationships	12	69 (84, 10-300)
Decision making for health in general	12	60 (73, 9-300)
Decision making for sexual health	12	57 (70, 10-270)
Communicating values about sex	10	53 (78, 5-300)
Identifying/avoiding risky sexual situations	12	56 (48, 5-180)
Teen pregnancy	12	61 (66, 10-270)
HIV/AIDS and other STIs	12	79 (77, 10-300)
Abstinence	12	94 (146, 10-600)
Condoms and/or contraception	11	76 (83, 6-270)
Media influence on sexual health	11	36 (46, 2-180)
Dating violence	10	39 (45, 5-180)
Total minutes implementing sexual health education	████████	643 (715, 100-3000)

[a] The average numbers of minutes refers only to the schools implementing that topic.

Table E.6. Percentage of IYG activities completed by lesson, 7th grade

Lesson – 7th grade	Number of activities	N	Minimum	Maximum	Mean	Standard deviation
1. It's Your Game: Pre-Game Show	7	92	71.4	100	96.0	8.0
2. Keeping It Real: Among Friends	7	92	85.7	100	99.7	2.1
3. Keeping It Real: Among Friends (computer lesson)	4	92	50.0	100	96.2	11.7
4. It's Your Game: Playing by Your Rules	6	88	66.7	100	97.9	7.1
5. It's Your Game: Playing by Your Rules (computer lesson)	4	92	75.0	100	98.4	6.2
6. Protecting Your Rules: A Clear No (role play lesson)	8	92	75.0	100	98.9	4.0
7. Protecting Your Rules: Alternative Actions (role play lesson)	7	92	85.7	100	99.8	1.5
8. Know Your Body (computer lesson)	3	92	66.7	100	99.6	3.5
9. Keeping It Real: For Yourself	6	92	66.7	100	98.2	6.7
10. Playing by Your Rules: Regarding Sex (computer lesson)	5	92	40.0	100	98.0	7.9
11. Protecting Your Rules: Regarding Sex (role play lesson)	6	92	83.3	100	99.6	2.4
12. It's Your Game: Post-Game Show	5	92	80.0	100	98.5	5.3

Note. There were a total of 68 activities, with an average 98.4% implementation rate.

Table E.7. Percentage of IYG activities completed by lesson, 8th grade

Lesson – 8th grade	Number of activities	N	Minimum	Maximum	Mean	Standard deviation
1. It's Your Game: Pre-Game Show	9	99	88.9	100.0	99.2	2.9
2. Keeping It Real: Consequences of Pregnancy	6	99	33.3	100.0	98.3	7.7
3. Keeping It Real: Consequences of HIV & Other Sexually Transmitted Infections (computer lesson)	4	98	50.0	100.0	96.7	9.9
4. Keeping It Real: Consequences Sexually Transmitted Infections	6	99	33.3	100.0	96.8	9.4
5. Keeping It Real: Risk Reduction Strategies (computer lesson)	4	98	75.0	100.0	99.2	4.3
6. Playing by Your Rules: A Review	6	98	83.3	100.0	99.1	3.7
7. Playing by Your Rules: A Review (computer lesson)	4	98	75.0	100.0	99.2	4.3
8. Keeping It Real: Healthy Dating Relationships	7	99	85.7	100.0	99.0	3.7
9. Keeping It Real: Healthy Dating Relationships (computer lesson)	4	98	50.0	100.0	99.0	6.1
10. Playing by Your Rules: Regarding Sex (role play lesson)	9	99	66.7	100.0	98.9	4.3
11. It's Your Game: Free Time (computer lesson)	4	99	75.0	100.0	98.7	5.5
12. It's Your Game: Post-Game Show	5	99	80.0	100.0	98.6	5.2

Note. There were a total of 68 lessons, with an average 98.6% implementation rate.

Table E.8. External events during study time frame

Continuous Conditions

1. Two intervention schools in one district and 3 intervention schools in a different district were providing stipends to facilitators for implementing IYG.

2. One intervention school in one district, 2 intervention schools in a second district, and 1 school in a third district were using the funds from the research study to support salaries. Teachers may not have been aware of this funding source.

During the study cohort's 7th grade year

3. As a part of Teen Pregnancy Prevention Month in May 2012, the former Miss SC visited an equal number of intervention and comparison schools to do a group presentation on following your dreams and achieving your goals.

4. One intervention school hosted a Christian rapper to perform during a school wide assembly where he facilitated an alter call (call for salvations) during the service. The ACLU sued the school district.

5. At one comparison school, the health teacher noted that "a lot" of male and female students "began to engage in sexual activities with multiple partners."

6. One comparison school teacher noted that students were aware of pregnancies at the high school and that at other schools some students had been raped and others had gotten STIs.

During the study cohort's 8th grade year

7. One comparison school teacher noted that students were pregnant in the school.

8. At one participating school, a school staff person (not involved in the project) was charged with criminal sexual conduct with a minor in January 2013.

Appendix F: Sensitivity analyses

Given the decision to return to a sub-sample of comparison schools to increase parent consent return rates and baseline participation in winter 2012, sensitivity analyses were conducted to examine the impact on all outcomes using the original fall sample only. Additionally, sensitivity analyses were conducted for the primary sexual initiation outcome without inclusion of the school-level proxy baseline covariate representing percent of students reporting they had sex at baseline given this variable was not a true baseline outcome variable and was not significantly imbalanced at $p<.15$ across conditions in multilevel screening tests. The results of these two sensitivity analyses are summarized in Table F.1. Additionally, layered analyses were conducted post-hoc to investigate the impact of gradual inclusions of covariates or sets of covariates on the negative finding in the benchmark analysis (see Table F.2).

Table F.1. Sensitivity of impact analyses using data from IYG student survey to address the primary research question

Intervention compared with comparison	Benchmark intervention percent	Benchmark comparison percent	Benchmark difference (p-value)	Sensitivity 1[a] intervention percent	Sensitivity 1[a] comparison percent	Sensitivity 1[a] difference (p-value)	Sensitivity 2[b] intervention percent	Sensitivity 2[b] comparison percent	Sensitivity 2[b] difference (p-value)
Initiation of sex (9th grade follow-up)	25.4	21.1	4.3 (0.039)	25.3	21.7	3.6 (0.171)	24.5	20.4	4.1 (0.039)

Table F.2. Sensitivity of impact analyses using data from IYG student survey to address the secondary research questions

Intervention compared with comparison	Benchmark intervention percent	Benchmark comparison percent	Benchmark difference (p-value)	Sensitivity 1[a] intervention percent	Sensitivity 1[a] comparison percent	Sensitivity 1[a] difference (p-value)	Sensitivity 2[b] intervention percent	Sensitivity 2[b] comparison percent	Sensitivity 2[b] difference (p-value)
Initiation of sex (8th grade follow-up)	14.9	12.9	2.0 (0.204)	14.7	13.5	1.2 (0.619)	14.5	12.7	1.8 (0.245)
Had sex in past 3 months (9th grade follow-up)	16.3	13.6	2.7 (0.135)	16.1	14.2	1.9 (0.355)	15.6	12.9	2.7 (0.115)
Had sex without effective birth control in the past 3 months (9th grade follow-up)	6.7	6.1	0.6 (0.586)	6.5	6.3	0.2 (0.893)	5.9	5.9	0.7 (0.499)

Table F.3. Layered analyses addressing the impact of sets of independent variables on results for the primary research question

Primary outcome measure	Intervention adjusted percent	Comparison adjusted percent	Percent difference (p-value of difference)	Odds ratio
1. Treatment	26.0	25.0	1.0 (0.751)	1.05
2. Treatment + **Design factors**[a]	25.2	24.2	1.0 (0.731)	1.06
3. Treatment + Design factors + **Late enrollment**[b]	25.9	23.2	2.7 (0.376)	1.16
4. Treatment + Design factors + Late enrollment + **Demographics**[c]	25.5	22.3	3.2 (0.269)	1.19
5. Treatment + Design factors + Late enrollment + Demographics + **Student-level covariates**[d]	25.4	21.7	3.7 (0.202)	1.23
6. Treatment + Design factors + Late enrollment + Demographics + Student-level covariates + **School-level exposure to 9th grade EBP**[e]	25.3	21.7	3.6 (0.171)	1.23
7. Treatment + Design factors + Late enrollment + Demographics + Student-level covariates + **School-level initiation rate**[f]	25.5	20.9	4.6 (0.053)	1.29
8. Treatment + Design factors + Late enrollment + Demographics + Student-level covariates + School-level exposure to 9th grade EBP + School-level initiation rate[g]	25.4	21.1	4.3 (0.039)	1.27

Source: Second follow-up survey (9th grade), Mar-Sept 2014, administered 12-18 months after the program.

[a]   Design factors include: (a) 7th grade school enrollment for study cohort, and (b) school configuration (MS only or MS/HS combined)

[b]   Late enrollment is an indicator of the sub-sample of students who completed the baseline survey approximately 3 months after the rest of the sample, in late January through mid-February 2012.

[c]   Demographics include the standard demographics of age, gender, and race/ethnicity.

[d]   Other student-level covariates include other *a priori* covariates that (a) the literature suggests are related to risky sexual behavior, (b) were missing for less than 10% of cases, and (c) were significantly different between conditions at baseline at $p < .15$ (accounting for clustering). Covariates included in the final models were the number of months since baseline and academic grades at baseline.

[e]   School-level exposure to 9th grade EBP refers to the value used for randomization of the expected exposure to an evidence-based TPP program in 9th grade.

[f]   School-level initiation rate refers to the school-level proxy baseline covariate representing percent of students reporting they had sex at baseline.

[g]   This analytic model is the same as the benchmark model presented in the Results section and in Table F.1. above.

Appendix G: Sample attrition analyses

Attrition analyses were conducted to assess which, if any, key sample characteristics (i.e., intervention condition, indicator of late enrollment, and all student-level covariates assessed for baseline equivalence) were associated with students not completing a second follow-up survey in the spring of 9th grade and whether the association(s) differed by intervention condition . (They did not assess the likelihood of being dropped from the final analytic models due to missing covariate values.) For attrition analyses, the primary sample consisted of students who reported never having had sex on the baseline survey (n=2,843); the secondary sample consisted of all students who completed baseline survey (n=3143). Analyses were adjusted for clustering and involved three steps. (1) A separate regression model was estimated in which each characteristic, as well as intervention condition indicator, served as the independent variable and an indicator of attrition was the dependent variable. The characteristic/intervention indicator was considered related to attrition if the p-value was less than or equal to .05, according to the Wald test. (2) For characteristics significantly associated with the attrition indicator in Step 1, a model was analyzed with the attrition indicator as the dependent variable and intervention condition, the significant characteristic, and an interaction term of the two were independent variables. Attrition was considered to be differential by intervention condition for a given characteristic if the p-value of the interaction term was less than or equal to .05, according to the Wald test. (3) Step 3 models would have been conducted for characteristics whose relation to attrition significantly differed by condition in Step 2. However, this step was not needed due to Step 2 results.

Table G.1. Baseline characteristics not related to student attrition in the primary sample

Baseline characteristic – primary sample	Overall p-value	Interpretation
Treatment arm (0=comparison, 1=intervention)	0.639	Characteristic not related to attrition.
Completed baseline survey in February 2012 (0=no, 1=yes)	0.593	Characteristic not related to attrition.
Gender (0=male, 1=female)	0.820	Characteristic not related to attrition.
Race/Ethnicity: Black (referent is White)	0.950	Characteristic not related to attrition.
Race/Ethnicity: Hispanic/Latino(a) (referent is White)	0.101	Characteristic not related to attrition.
Race/Ethnicity: Other race (referent is White)	0.055	Characteristic not related to attrition.
Primary language at home is English (0=no, 1=yes)	0.911	Characteristic not related to attrition.
Live at other home some of the time (0=no, 1=yes)	0.070	Characteristic not related to attrition.
Father's level of education (1-6, higher value indicates more education)	0.356	Characteristic not related to attrition.

Baseline characteristic – primary sample	Overall p-value	Interpretation
Importance of faith (1-4, higher value indicates greater importance)	0.081	Characteristic not related to attrition.

Table G.2. Baseline characteristics related to student attrition in the primary sample, but which do not differ by condition

Baseline characteristic – primary sample	Overall p-value	Overall odds ratio	p-value for difference in OR by condition	Interpretation
Age at baseline (in years)	0.000	1.51	0.743	Older students were more likely to attrit, but this pattern didn't differ by condition.
Number of biological parents in the home (0-2)	0.000	0.60	0.591	Students with fewer parents in the home were more likely to attrit, but this patter didn't differ by condition.
Mother's level of education (1-6, higher value indicates more education)	0.000	0.89	0.599	Students with less educated mothers were more likely to attrit, but this pattern didn't differ by condition.
Mom was a teen parent (0=no, 1=yes)	0.000	1.77	0.281	Students whose mothers were teen parents were more likely to attrit, but this pattern didn't differ by condition.
Academic grades (1-4, higher value indicates better grades)	0.000	0.66	0.557	Students with poorer academic grades were more likely to attrit, but this pattern didn't differ by condition.
Educational aspirations (1-6, higher value indicates plan to go further in school)	0.010	0.91	0.686	Students with lower academic aspirations were more likely to attrit, but this pattern didn't differ by condition.
Frequency of attending religious services (1-4, higher value indicates greater frequency)	0.016	0.93	0.592	Students who attended religious services less often were more likely to attrit, but this pattern didn't differ by condition.
Number of days in last 30 had 1 or more alcoholic drinks (1-7, ordinal, higher value indicates more days)	0.020	1.16	0.354	The more days (in the 30 days prior to baseline) students reported drinking one or more alcoholic drinks, the more likely they were to attrit, but this pattern did not differ by condition.
Honesty with which answered baseline survey (1-5, higher value indicates less honesty)	0.034	1.15	0.195	The less honest students were answering the survey, the more likely they were to attrit, but this pattern didn't differ by condition.

Table G.3. Baseline characteristics not related to student attrition in the secondary sample

Baseline characteristic – secondary sample	Overall *p*-value	Interpretation
Treatment arm (0=comparison, 1=intervention)	0.938	Characteristic not related to attrition.
Completed baseline survey in February 2012 (0=no, 1=yes)	0.738	Characteristic not related to attrition.
Gender (0=male, 1=female)	0.342	Characteristic not related to attrition.
Primary language at home is English (0=no, 1=yes)	0.547	Characteristic not related to attrition.
Father's level of education (1-6, higher value indicates more education)	0.070	Characteristic not related to attrition.

Table G.4. Baseline characteristics related to student attrition in the secondary sample, but which do not differ by condition

Baseline characteristic – secondary sample	Overall *p*-value	Overall odds ratio	*p*-value for difference in OR by condition	Interpretation
Age at baseline (in years)	0.000	1.81	0.314	Older students were more likely to attrit, but this pattern didn't differ by condition.
Race/Ethnicity: Black (referent is White)	0.433	NA	NA	Black students were more likely to attrit than whites, but this pattern didn't differ by condition.
Race/Ethnicity: Hispanic/Latino(a) (referent is White)	0.122	NA	NA	Latino students were marginally more likely to attrit that whites, but this pattern didn't differ by condition.
Race/Ethnicity: Other race (referent is White)	0.012	1.44	0.922	Students of other races were more likely to attrit than whites, but this pattern didn't differ by condition.
Live at other home some of the time (0=no, 1=yes)	0.006	1.30	0.448	Students who lived in another home some of the time were more likely to attrit, but this pattern didn't differ by condition.
Number of biological parents in the home (0-2)	0.000	0.58	0.978	Students with fewer parents in the home were more likely to attrit, but this pattern didn't differ by condition.
Mother's level of education (1-6, higher value indicates more education)	0.000	0.86	0.774	Students with less educated mothers were more likely to attrit, but this pattern didn't differ by condition.

Baseline characteristic – secondary sample	Overall p-value	Overall odds ratio	p-value for difference in OR by condition	Interpretation
Mom was a teen parent (0=no, 1=yes)	0.000	1.91	0.253	Students whose mothers were teen parents were more likely to attrit, but this pattern didn't differ by condition.
Academic grades (1-4, higher value indicates better grades)	0.000	0.64	0.901	Students with poorer academic grades were more likely to attrit, but this pattern didn't differ by condition.
Educational aspirations (1-6, higher value indicates plan to go further in school)	0.000	0.88	0.642	Students with lower academic aspirations were more likely to attrit, but this pattern didn't differ by condition.
Importance of faith (1-4, higher value indicates greater importance)	0.016	0.88	0.430	Students' whose faith was less important to them were more likely to attrit, but this pattern didn't differ by condition.
Frequency of attending religious services (1-4, higher value indicates greater frequency)	0.002	0.92	0.830	Students who attended religious services less often were more likely to attrit, but this pattern didn't differ by condition.
Number of days in last 30 had 1 or more alcoholic drinks (1-7, ordinal, higher value indicates more days)	0.005	1.17	0.904	The more days (in the 30 days prior to baseline) students reported drinking one or more alcoholic drinks, the more likely they were to attrit, but this pattern didn't differ by condition.
Honesty with which answered baseline survey (1-5, higher value indicates less honesty)	0.001	1.28	0.850	The less honest students were answering the survey, the more likely they were to attrit, but this pattern didn't differ by condition.
Ever had sex (0=no, 1=yes)	0.000	2.54	0.05	Students who had ever had sex were more likely to attrit, but this pattern didn't differ by condition.
Had sex in the past 3 months (0=no, 1=yes)	0.000	2.96	0.218	Students who had sex in the 3 months prior to baseline were more likely to attrit, but this pattern didn't differ by condition.
Had sex without effective birth control in past 3 months (0=no, 1=yes)	0.000	4.01	0.123	Students who had sex without effective birth control in the 3 months prior to baseline were more likely to attrit, but this pattern didn't differ by condition.