

**Findings from the
Replication of an
Evidence-Based
Teen Pregnancy
Prevention
Program**

Evaluation of It's Your Game...Keep It Real in Houston, TX

Final Impact Report for

University of Texas Health Science Center - Houston

Revised: July 14, 2016

Prepared by

ETR

Karin Coyle, PhD, Pamela Anderson, PhD, BA Laris, MPH, Tracy Unti, BS,
Heather Franks, MA, and Jill Glassman, PhD, MSW

Recommended Citation

Coyle, K., Anderson, P., Laris, BA, Unti, T., Franks, H., & Glassman, J. (2016). Evaluation of It's Your Game...Keep It Real In Houston, TX: Final report. Scotts Valley, CA: ETR Associates.

Acknowledgements:

The authors gratefully acknowledge the contributions to this study by the program, evaluation, and data collector team members who helped ensure all aspects of the project were implemented successfully. Additionally, we express our sincere gratitude to the district representatives, principals, teachers, school staff, and students who participated in the project. The program and evaluation teams are also grateful for the ongoing reinforcement from Cornerstone Consulting, who provided instrumental support to school sites throughout the project. Finally, the evaluation team would like to express our appreciation to Dr. Elizabeth McDade-Montez and Susan Potter for their support in report preparation.

This publication was prepared under Grant Number TP1AH000072 from the Office of Adolescent Health, U.S. Department of Health & Human Services (HHS). The views expressed in this report are those of the authors and do not necessarily represent the policies of HHS or the Office of Adolescent Health.

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

I. Introduction

A. Introduction and study overview

Teen pregnancy and childbearing are serious public health issues in Texas, particularly in Harris County, the community targeted for this initiative. Texas has the 5th highest teen birth rate among females aged 15–19 years in the nation (41 per 1,000 in Texas vs. 26.5 per 1,000 in the U.S.), and the 2nd highest among school-aged females (aged 15–17 years) (Martin et al., 2015). Moreover, Texas has the highest repeat teen birth rate in the nation (21% in Texas vs. 17% in the U.S.) (National Campaign to Prevent Teen and Unplanned Pregnancy, 2015). There are racial/ethnic disparities in the Texas teen birth rate; the teen birth rate (i.e., 15-19 year olds) among Hispanics (62 per 1,000) and blacks (44.1 per 1,000) is significantly higher than that among non-Hispanic whites (26.3 per 1,000) (Ventura et al., 2014).

In Harris County the teen birth rate in 2012 (most recent year for which data were available) was 42.7 per 1,000 (Texas Department of State Health Services, 2012a), which surpasses the U.S. teen birth rate in 2012 (Martin et al., 2013) and 2013 (Martin et al., 2015). Teen births among minority youth are higher in Harris County, compared with state rates. About 70% of teen births in Harris County occur among Hispanic females, while 21% occur among Black females (Harris County Healthcare Alliance, 2012; Texas Department of State Health Services, 2012a). Data from the 2013 Youth Risk Behavior Survey (YRBS) of high school students indicate that Houston youth are more likely than other U.S. youth to engage in many sexual risk-taking behaviors: they are more likely to have sex before age 13 (7.9% vs. 5.6%), less likely to use birth control (13.9% vs. 25.3%), and less likely to use any method to prevent pregnancy at last sex (75.8% vs 86.3%) (CDC, 2015). Houston high school students are also less likely than U.S.

Final: 7.14.16

youth to report having received any sexual health education in school (68.3% vs. 85.3%) (CDC, 2015). Given the various health challenges confronting youth in Harris County, it is critical that school districts employ teen pregnancy prevention programs that are evidence-based. In response to the Teen Pregnancy Prevention (TPP) Tier 1 funding announcement from the Office of Adolescent Health, The University of Texas Health Science Center at Houston partnered with local school districts in Harris County to train teachers to implement *It's Your Game...Keep It Real!* (IYG) an evidence-based TPP program designed and tested with urban youth.

IYG has been tested in two separate studies (Tortolero et al., 2010; Markham 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 final 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 HHS 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 final 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 HHS 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 Harris County, Texas middle schools funded through a grant from the Office of Adolescent Health to the University of Texas Health Science Center at Houston. ETR was contracted to conduct the evaluation. This report adds to the literature on replicating evidence-based programs under different conditions.

B. Primary research question(s)

The primary research question addressed overall program impact on the combined outcome of vaginal or oral sexual initiation: *What is the impact of the IYG program relative to the usual health curriculum on initiation of either vaginal or oral sex by the end of 9th grade (approximately one year after the end of the program) for students reporting “no” to ever had vaginal or oral sex at baseline?*

C. Secondary research question(s)

Secondary behavioral outcomes addressed overall program impact on either vaginal or oral sexual initiation: (1) *What is the impact of the IYG program relative to the usual health curriculum on initiation of vaginal intercourse by the end of 9th grade for students reporting “no” to ever had vaginal intercourse at baseline?* And (2) *What is the impact of the IYG program relative to the usual health curriculum on initiation of oral sex by the end of 9th grade for students reporting “no” to ever had oral sex at baseline?*

II. Program and comparison programming

A. Description of program as intended

IYG 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 the program’s underlying behavior change theories--social cognitive theory, social influence models, and the theory of triadic influence—which represent an array of factors (e.g., environmental, personal, social) that influence behavior (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, (5) planning for adoption, implementation and sustainability, and (6) planning for 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. 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). 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, IYG served as the primary source for reproductive health content in the 10 intervention schools.

B. Description of counterfactual condition

Each school in the comparison condition provided its usual health and sex education program, which varied by district because sexual health and HIV education are not mandated in Texas. Schools were not considered eligible for participation in the study if an evidence-based 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. As part of the evaluation, data were collected from the comparison school health teachers about use of existing programs; these teachers confirmed that

they did not use evidence-based or promising programs to teach about sexual health during the period the intervention schools were teaching the 7th and 8th graders enrolled in the study.

III. Study design

A. Sample recruitment

1. School sample

The study involved working with selected school districts and schools in Harris County, Texas. The University of Texas Health Science Center at Houston (UTHSC) recruited schools via school district administrators during 2010–2011, the year prior to commencing the evaluation. Eligibility was determined at three levels: *district, school, and student*. Ten districts representing 73 middle schools were screened for participation. Participating *districts* had to meet the following criteria:

- Contain 2 or more schools with 7th and 8th grades;
- Provide a list of schools that would be willing to participate and agree to the conditions of the study if eligible.

Invitation letters were then sent to these middle schools with 7th and 8th grades in Harris County. Participating *schools* had to meet the following criteria:

- Not currently implementing IYG or using another evidence-based sex education program in 7th and 8th grades;
- Have 7th grade enrollment of more than 150 students;
- Have no known implementation, logistical, or cooperation issues that would make participation difficult.

To gather more information about known implementation, logistical or cooperation issues, UTHSC conducted suitability assessments for each school. These assessments involved

Final: 7.14.16

conversations with district coordinators, school staff, training staff, and staff from other projects that previously worked in the districts and schools. The assessments ascertained for each school were based on:

- Previous cooperation in all evaluation activities, including data collection and recruitment activities for other projects;
- Logistical capacity to implement the evaluation, including space for data collection, scheduling flexibility for evaluation activities;
- Feasibility of fidelity to the IYG implementation plans as submitted by schools and approved by UTHSC.

Schools were further excluded where there was a concern in any of the areas above, and where implementation plans were unclear.

Of the 73 schools screened within the 10 districts of Harris County, 20 middle schools within 5 school districts met the eligibility criteria and agreed to participate. Ten schools were randomized to be in the treatment group (received the IYG curriculum) and 10 were randomized to be in the comparison group (continued to receive their regular school-based health education program); the comparison schools were informed that they could implement IYG after the final data collection was complete.

All 20 schools are urban middle schools across Harris County, Texas with total enrollments ranging from 500 to 1,950 students at the time of randomization. Harris County represents one of the most diverse and disadvantaged counties in the nation: 38% of residents are Hispanic, 20% are African American, one-third of adults speak a language other than English, and over 23% of Harris county children live in poverty. At the time of randomization, the percent

of students who qualified for free lunch across the 20 participating schools was 79%, ranging from 47% to 90%.

2. Youth sample

Youth were eligible to participate if they were enrolled in 7th grade at a participating school in fall 2012, 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., positive permission by a parent/legal guardian) was obtained prior to data collection and at one time for all study activities. We used a mix of census and sampling when recruiting for participation. In schools with 250 or fewer 7th grade students, we distributed consent forms to all students using the process described below. For schools that have 7th grade enrollments of greater than 250 students, we sampled classes to achieve a starting cohort size of 180 students (that is, 180 students that received consent forms). The study includes one cohort and follows them from 7th grade through 9th grade.

ETR data collection staff, blind to school status and student participation in the intervention, visited each 7th grade class across the 20 middle schools and presented information about the study to the students. The presentation described the purpose, general design, and enrollment criteria to eligible youth during classroom time. Parent consent forms and information about the study were sent home to parents via their children. Parents were asked to return their child's consent form to their classroom teacher by a designated date.

Parent consent return was promoted throughout the baseline data collection period. For example, ETR staff checked in with designated teachers to obtain signed consent forms on repeat visits to the school prior to data collection. To help ensure that signed parental consent forms were returned, a \$25 stipend was given to teachers for each class that returned parental consent

forms for 90% of students, regardless of whether parents agreed to allow students to participate. 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. Finally, to encourage the timely return of consent forms at each study school, all students from each school who returned their consent forms by a set date (within 10 days of receiving them) were entered into a school-level drawing to receive an iPod Touch; 3 iPod Touch devices were distributed per school. If students forgot to bring back their consent forms after multiple reminders over a period of approximately 2 school weeks, trained ETR staff used a scripted protocol to obtain verbal consent from students’ parents or legal guardians via telephone. Student assent was obtained from all students with parental consent immediately prior to administering the survey.

Overall, 93% of the 3,565 eligible students returned parent consent forms; 73% had positive parent consent. In total, 67.4% of eligible students (n = 2,403) completed a baseline survey between September 2012 and early March 2013, representing 67.8% of eligible youth in intervention schools (n = 1,232) and 67.0% of eligible youth in comparison schools (n = 1,171).

B. Research design

The study involved an experimental group-randomized trial design in which the 20 participating schools were randomized to receive IYG (intervention condition) or serve as comparison sites. Randomization was performed at the school level in fall 2011, one year prior to baseline data collection by a previous evaluator¹ using a multi-attribute randomization protocol (Graham et al., 1984) to optimize the balance of the following variables across study conditions: 7th-grade enrollment in the school; percent of Black students in the school; percent

¹ Intervention schools began implementing IYG with a cohort of 7th-grade students in fall 2011 following randomization. However, due to initial complications with the evaluation the baseline was delayed one year, and the fall 2011 cohort was not consented or surveyed and, ultimately, not included in the study. The evaluation examined impacts for a cohort of students who were in the 7th grade in fall 2012.

Final: 7.14.16

of Hispanic students in the school; percent of students in the school who receive a free lunch. Specifically, these four variables were combined into a single index using principal components analysis. Within each district, schools whose index scores were closest to each other were paired. Finally, within each pair one school was randomly assigned to the intervention condition and the other to the comparison condition. The school district was used as a stratification variable to balance the number of schools within a district assigned to each condition. There were five school districts and an even number of schools within each district.

School administrators were notified of their condition after randomization but before baseline data collection. The first evaluator planned to start baseline data collection in fall 2011, but, due to logistical challenges faced by the evaluator, was unable to start baseline data collection. All 20 schools were informed that the baseline survey would be postponed for a year. Schools in the intervention condition were encouraged to use the school year as a pilot opportunity, allowing 7th grade teachers to practice using IYG. All 10 intervention schools completed implementation plans to pilot IYG; no data were collected on how many actually taught IYG that year. After a change in evaluator in 2012, the baseline survey was rescheduled for fall 2012. Because so few consent forms were distributed in fall 2011, few parents and students were aware of intervention status. Administrators at all 20 schools and 7th grade teachers were informed of the change in study timing.

Once the fall 2012 baseline was underway, students included in the evaluation were surveyed prior to the program (Fall-Winter 2012-2013), approximately 3 months after the end of the 8th grade intervention (Winter-Spring 2014), and one year later when they were in 9th grade (Winter-Spring 2015). Consent and assent procedures occurred at baseline only and covered all three data collection time points.

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; the survey also assessed student demographics and other background characteristics. At each time point, students in matched intervention and comparison schools were surveyed during the same time frame (within two weeks of one another) and the condition surveyed first in a pair was varied systematically across the pairs. See Appendix A for specifics on data collection timeframes.

Data were collected by trained data collectors in school using audio-enhanced computer assisted surveys via laptops through the study schools. At baseline, students with consent and assent received a \$10 gift card for completing the survey; they also were allowed to keep the headsets they used to complete the survey. At the first follow-up (spring of 8th grade), students received a \$10 gift card and could keep the headsets. At the second follow-up (spring of 9th grade), students received a \$15 gift card for their participation and could keep the headsets. At each follow-up, 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). As with students surveyed in school, students who completed a follow-up survey online, by mail, or by phone received 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 each year of programming. Implementation logs, created by IYG developers to measure program adherence 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 5 school days of teaching an IYG lesson. Observations conducted by trained evaluation and program staff assessed both adherence and quality of implementation; 3% 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) covering a different lesson in each observation. Ultimately, however, the observation sample was one of convenience due to teachers' 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 classroom of students. Incentives were tied to attendance submission in combination with the implementation logs. Implementation log and observation data were reviewed by IYG project staff on a weekly basis, allowing them to provide ongoing technical assistance (TA) to facilitators as needed. Project staff notified district-level coordinators when issues were raised in their schools. Project staff noted suggestions and challenges and discussed in the district coordinator meetings. IYG facilitators completed online reaction surveys at the end of each school year in which they provided information about their training, background, and experience with IYG implementation. Health educators at comparison middle schools also completed online surveys asking about the content of and time spent implementing any sexual health education to the study cohort. IYG facilitators and comparison school health educators received \$20 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 and secondary behavioral outcomes are described in Tables III.1 and III.2, respectively.

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

Outcome name	Description of outcome	Timing of measure relative to program
Initiation of vaginal or oral sex	<p>The variable is a yes/no measure of whether a person has ever had vaginal OR oral sex. The measure is created from the following items on the survey:</p> <ul style="list-style-type: none"> • “Have you ever had sexual intercourse?” (defined in survey as penis in vagina) • “Have you ever had oral sex?” <p>Participants who respond yes they have had sexual intercourse OR yes they have had oral sex are coded as 1 for yes; those who respond no they have not had sexual intercourse AND no they have not had oral sex are coded as 0 for no.</p>	12 months after program ends (spring of 9th grade)

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 vaginal intercourse	<p>The variable is a yes/no measure of whether a person has ever had vaginal intercourse. The measure is based on the following item on the survey:</p> <ul style="list-style-type: none"> • “Have you ever had sexual intercourse?” (defined in survey as penis in vagina) <p>Respondents who respond yes they have had sexual intercourse are coded as 1 for yes and those who respond no they have not had vaginal intercourse are coded as 0 for no.</p>	12 months after program ends (spring of 9th grade)
Initiation of oral sex	<p>The variable is a yes/no measure of whether a person has ever had oral sex. The measure is based on the following item on the survey:</p> <ul style="list-style-type: none"> • “Have you ever had oral sex?” <p>Respondents who respond yes they have had oral sex are coded as 1 for yes and those who respond no they have not had oral sex are coded as 0 for no.</p>	12 months after program ends (spring of 9th grade)

E. Study sample

Twenty schools were recruited into the study; all 20 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. At the time of baseline data collection, 3,565 eligible

Final: 7.14.16

students were enrolled in sampled classes at participating schools—1,818 students at intervention schools and 1,747 at comparison schools. The final baseline sample consisted of 2,403 youth, 1,232 students at intervention schools and 1,171 students at comparison schools; 67.8% and 67.0% of eligible youth, respectively, for an overall participation rate of 67.4%. Of the students taking a baseline survey, 82.5% of intervention and 85.3% of comparison students completed the first follow-up survey approximately 3 months after the program in the spring of their 8th grade year (55.9% and 57.2 % of those eligible, respectively), and 78.1% of intervention and 81.1% of comparison students completed the final follow-up survey 12 months post-program in the spring of their 9th grade year (representing 52.9% of eligible intervention students and 54.4% of eligible comparison students). More details regarding the study sample are included in Appendix C.

By definition, the primary sample was comprised only of those who reported not having had vaginal or oral sex at baseline—1,069 intervention students and 947 comparison students. Of these, 806 intervention and 740 comparison students completed a final follow-up survey in spring of their 9th grade, which represents 75.4% and 78.1% of intervention and comparison students reporting no vaginal or oral sex at baseline. After removing missing values for covariates included in analysis models, the final primary analytic sample included 804 intervention and 739 comparison students. The secondary analytic samples varied by outcome. For the outcome focused on initiation of vaginal intercourse, the sample included 1,582 students—820 in the intervention condition and 762 in the comparison condition. The sample for the outcome focused on initiation of oral sex included 1,599 students—829 intervention and 770 comparison. Tables III.3 to III.5 provide more details regarding key demographic characteristics

on the samples used to assess initiation of vaginal or oral sex, vaginal sex only and oral sex only, respectively.

F. Baseline equivalence

The following variables were assessed for equivalence between the intervention and comparison conditions at baseline for the primary and secondary behavioral outcomes because literature indicates they are typically related to risky sexual behavior (Kirby & Lepore, 2007): age, gender, race/ethnicity, number of years living in the U.S., academic grades, two indicators of religiosity, two indicators of home structure, and maternal history of teen parenthood. The intervention and comparison groups were identical on the baseline measure of the sexual initiation outcomes at baseline because, by definition, students included in these analyses had not yet initiated the behaviors. Multilevel regression analyses were conducted with the variable of interest as the dependent variable and the intervention indicator as the independent variable. 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.5 shows results of these baseline equivalence analyses for the primary and secondary analytic samples. Results showed there were no statistically significant differences between the intervention and comparison group means at baseline on any of the assessed variables, suggesting that at the start of the study the groups were equivalent on these measures.

Table III.3. Summary statistics of key baseline measures for youth completing IYG student survey for primary analytic sample filtered by no vaginal or oral sex initiation.

Baseline measure	Intervention % or mean (standard deviation)	Comparison % or mean (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison p-value of difference ^a
Age	13.03 (.57)	12.95 (.56)	0.08	0.252
Gender (female)	55.0%	54.4%	0.6	0.871
Race/ethnicity: Black	24.4%	24.5%	-0.1	0.833

Baseline measure	Intervention % or mean (standard deviation)	Comparison % or mean (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison p -value of difference ^a
Race/ethnicity: Hispanic	63.1%	63.3%	-0.2	0.892
Race/ethnicity: Other	12.4%	12.2%	0.2	0.969
Years in USA	11.67 (2.16)	11.54 (2.31)	0.13	0.488
Grades (4 = mostly A's & B's, 1 = mostly D's & F's)	3.46 (.66)	3.39 (.67)	0.07	0.224
Importance of religion (1 = not important, 4 = very important)	2.93 (.83)	2.95 (.86)	-0.02	0.883
Frequency of religious services (1 = never, 6 = more than once per week)	4.03 (1.52)	3.89 (1.61)	0.14	0.145
Number of parents in household (0-2 parents)	1.58 (.57)	1.58 (.60)	0	0.993
Live in multiple homes	18.5%	18.8%	-0.3	0.945
Mom was a teen parent	28.7%	31.1%	-2.4	0.395
School-level rate of 7 th graders reporting ever had sex	7.02%	12.16%	-5.14%	<.001 ^c
Sample size ^b	804	739		

^aThe p -values are adjusted for clustering at the level of random assignment.

^b The primary analytic sample was comprised of students who completed a baseline survey, a follow-up survey, provided values for covariates included in the final analysis models, and reported not having had vaginal or oral sex at baseline.

^c p -value not adjusted for clustering since this is a school-level variable

Table III.4. Summary statistics of key baseline measures for youth completing IYG student survey for secondary analytic sample filtered by no vaginal sex initiation.

Baseline measure	Intervention % or mean (standard deviation)	Comparison % or mean (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison p -value of difference ^a
Age	13.02 (.58)	12.96 (.56)	0.06	0.367
Gender (female)	54.6%	53.7%	0.9	0.792
Race/ethnicity: Black	24.9%	24.7%	0.2	0.869
Race/ethnicity: Hispanic	62.3%	62.7%	-0.4	0.927
Race/ethnicity: Other	12.7%	12.6%	0.1	0.993
Years in USA	11.69 (2.14)	11.54 (2.31)	0.15	0.396
Grades (4 = mostly A's & B's, 1 = mostly D's & F's)	3.46 (.66)	3.39 (.67)	0.07	0.212
Importance of religion (1 = not important, 4 = very important)	2.93 (.84)	2.96 (.86)	-0.03	0.720
Frequency of religious services (1 = never, 6 = more than once per week)	4.00 (1.53)	3.89 (1.61)	0.11	0.273
Number of parents in household (0-2 parents)	1.57 (.57)	1.57 (.60)	0	0.910
Live in multiple homes	18.4%	19.2%	-0.8	0.760
Mom was a teen parent	28.7%	31.5%	-2.8	0.307
School-level rate of 7 th graders reporting ever had sex	7.08%	12.28%	-5.20%	<.001 ^c
Sample size ^b	820	762		

^aThe p -values are adjusted for clustering at the level of random assignment.

^bThe secondary analytic sample was comprised of students who completed a baseline survey, a follow-up survey, provided values for covariates included in the final analysis models, and reported not having had vaginal sex at baseline.

^c p -value not adjusted for clustering since this is a school-level variable

Table III.5. Summary statistics of key baseline measures for youth completing IYG student survey for analytic sample filtered by no oral sex initiation.

Baseline measure	Intervention % or mean (standard deviation)	Comparison % or mean (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison p -value of difference ^a
Age	13.02 (.57)	12.97 (.57)	0.05	0.564
Gender (female)	54.2%	54.0%	0.2	0.991
Race/ethnicity: Black	25.1%	24.8%	0.3	0.860
Race/ethnicity: Hispanic	62.6%	62.6%	0	0.865
Race/ethnicity: Other	12.2%	12.6%	-0.4	0.777
Years in USA	11.69 (2.13)	11.56 (2.31)	0.13	0.428
Grades (4 = mostly A's & B's, 1 = mostly D's & F's)	3.46 (.66)	3.37 (.68)	0.09	0.136
Importance of religion (1 = not important, 4 = very important)	2.94 (.85)	2.96 (.86)	-0.02	0.944
Frequency of religious services (1 = never, 6 = more than once per week)	4.02 (1.53)	3.87 (1.61)	0.15	0.094
Number of parents in household (0-2 parents)	1.57 (.58)	1.56 (.60)	0.01	0.877
Live in multiple homes	18.7%	19.4%	-0.7	0.779
Mom was a teen parent	28.5%	31.9%	-3.4	0.228
School-level rate of 7 th graders reporting ever had sex	7.07%	12.22%	-5.15%	<.001 ^c
Sample size ^b	829	770		

^aThe p -values are adjusted for clustering at the level of random assignment.

^bThe secondary analytic sample was comprised of students who completed a baseline survey, a follow-up survey, provided values for covariates included in the final analysis models, and reported not having had oral sex at baseline.

^c p -value not adjusted for clustering since this is a school-level variable

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. 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 vaginal or oral sex).

Each model included an indicator variable denoting intervention condition, age, gender and race/ethnicity measured at baseline, and a set of *a priori* identified outcome-related covariates measured at baseline. Outcome-related covariates were included only if they differed at $p < .15$ ² between the conditions in the appropriately filtered baseline sample of students who had a final follow-up; p-values reflect adjustment for clustering. Additionally, three of the variables used in the randomization process (district, 7th grade school enrollment, and percent of students in the school who received a free lunch) were included as covariates in the model, regardless of whether they were imbalanced at baseline; school-level race/ethnicity, which was used in randomization, was not included because it was already represented by individual-level demographic variables (race/ethnicity) and model parsimony was desired. Finally, an indicator representing the percentage of entering 7th grade students in the school reporting they ever had vaginal or oral sex at baseline was included. The latter was included in an attempt to control for

² Our covariate screening is derived from those suggested in Altman (1991) and Hosmer and Lemeshow (1989). In the latter, it is suggested that $p < .25$ as a screening criterion may be more appropriate than $p < .05$ because the latter often fails to identify variables that may be important to control. We “split the difference” and selected a $p < .15$ to preserve degrees of freedom of the model. We routinely include baseline outcome regardless of screening criteria (Pocock et al., 2002).

potential environmental or normative influences that may have resulted from the unexpectedly large observed imbalance in rates of reported vaginal or oral sex in the present study's sample of all entering 7th grade students taking a baseline survey (7.5% in the intervention condition and 12.5% in the comparison condition). This combined school-level indicator of the rates of sexual activity for entering 7th graders was used in the models for all 3 behavioral outcomes.

Missing data on baseline demographic covariates (e.g., age, race/ethnicity) were filled in when possible based on responses to relevant items on subsequent follow-up surveys. Missing data on the "ever had sex" item were recoded to 1 ("yes") if at least 3 other responses to secondary sexual behavior items indicated the student had sex. Inconsistent cases (e.g., "yes" to ever had sex at baseline and "no" at follow-up) were coded to missing.

One sensitivity analysis was conducted to understand the influence of including the covariate representing the percent of students reporting they ever had vaginal sex or oral sex at baseline. All analyses were conducted using STATA 13.1, which utilizes maximum likelihood methods for fitting multilevel models.

2. Implementation evaluation

Implementation data were analyzed using descriptive statistics and qualitative analysis. Results for analysis of adherence, quality, counterfactual experiences, and context are presented as frequency counts, percentages, averages, standard deviations, and/or ranges. See Appendix D for more detail on the implementation evaluation methods.

IV. Study findings

A. Implementation study findings

1. Adherence

Sessions delivered. Across 7th grade classes with complete log data (126 classes, or approximately 87% of classes implemented), facilitators delivered 10.4 out of 12 sessions

Final: 7.14.16

(86.7% of the curriculum) on average. Across 8th grade classes with complete log data (133 classes, or approximately 98% of classes implemented), facilitators delivered 11 of 12 sessions (97.4% of the curriculum) on average. IYG lessons are designed to be 50 minutes each for a total curriculum time of 600 minutes. The average duration of each session was 45 minutes during 7th grade and 45 minutes during 8th grade, equating to an average total of 468 minutes and 495 minutes of programming in each grade, respectively. Individual teachers were allowed to determine the frequency with which sessions were delivered. Forty five percent of the lessons were implemented daily, 13% every other day, 17% every 3 or 4 days, and 26% weekly or less than once a week by 7th grade teachers. Eighth grade teachers implemented 36% of lessons daily, 6% every other day, 33% every 3 or 4 days, and 25% weekly or less than once a week, on average.

Dosage received. In 7th grade, based on the attendance data received, students attended an average of 10.9 sessions (90.8% of 12 lessons), and 0 students did not attend any sessions. In 8th grade, students attended an average of 10.8 sessions, or 90.0% of 12 lessons, and 1 student, or 0.1%, did not attend any sessions. A substantial number of teachers did not report attendance (51% missing for 7th grade and 30% missing at 8th grade) despite incentives and repeat reminders. The lack of complete attendance data makes it difficult to fully assess program adherence related to student dosage in 7th and 8th grade.

Content covered. Teachers delivered an average of 95.3% of the IYG activities within 7th grade lessons (64 of 70 possible activities) and 93% in 8th grade lessons (62 of 67 possible activities).

Program staff and training. Thirty-seven IYG facilitators completed training in IYG and thereby met the qualifications to teach the curriculum. Twenty-seven facilitators implemented

Final: 7.14.16

7th grade IYG and 26 implemented it in the 8th grade. During 7th grade, the number of IYG facilitators per school ranged from one to four. During 8th grade, the number of IYG facilitators ranged from one to five. Most 7th and 8th grade facilitators were physical education (PE) teachers or PE and health teachers (85% and 90%, respectively); the others were contracted specifically as IYG facilitators. All teachers (100%) had access to support and technical assistance (TA) through their District Coordinators. See Table F.1. in Appendix F for more details on adherence.

2. Quality

During the 7th grade, four outside raters observed 52 IYG classes using a 5-point scale where 1 = poor, 3 = average, and 5 = excellent. In the first year, all observations were double coded and had high levels of interrater reliability. On ratings of teacher comfort level discussing sex-related topics, 47 observations (90.4%) included ratings for this indicator (a small number of teachers did not explicitly talk about sex-related topics). Of those that did receive a rating, the average rating was 4.6; 90.5% had a score of 4 or 5. On ratings of teacher rapport with students, the average rating was 4.5; 86.5% of the 7th grade ratings were a 4 or 5. Among observations during which students asked questions (90.4% of observations), 93.4% included ratings of a 4 or 5 on teachers' ability to address student questions (mean rating = 4.6). Observers also rated levels of youth engagement during 49 IYG sessions using the same 5-point scale. The average rating was 4.4; 84.3% had a score of 4 or 5.

In 8th grade, 47 observations were conducted by two outside raters using the same 5-point scale as described for the 7th grade. On ratings of teacher comfort level discussing sex-related topics, 34 observations (72.3%) included ratings for this indicator (27.7% of teachers did not explicitly talk about sex-related topics). Of those that did receive a rating, 94.1% had a score of 4

Final: 7.14.16

or 5 (mean rating = 4.9). On ratings of teacher rapport with students, the average rating was 4.7; 93.5% of the 8th grade ratings were a 4 or 5. Among observations during which students asked questions (59.6% of observations), 96.5% included ratings of a 4 or 5 on teachers' ability to address student questions (average rating was 4.9). Observers also rated levels of youth engagement during 47 IYG session observation; 72.4% had a score of 4 or 5 (average rating = 4.3).

Data collected from teacher and student reaction forms suggest that the program was well received. For example, 100% of 7th grade teachers reported on a teacher survey that they agreed or strongly agreed that teaching IYG was enjoyable (mean score = 3.25 on a scale of 1 = strongly disagree to 4 = strongly agree); further, 90% said they wanted to continue teaching IYG. Similarly, 90% of 8th grade teachers reported that they agreed or strongly agreed that teaching IYG is enjoyable (mean score = 3.24 on a scale of 1 = strongly disagree to 4 = strongly agree), and 81% reported wanting to continue teaching IYG (data not shown).

During spring 2014, after completing IYG lessons, over 20,000 students in 60 middle schools—50 involved in a larger dissemination effort supported by OAH and the 10 intervention middle schools in this replication evaluation—completed a paper-pencil survey about their experience in IYG. About 52% of these students were 7th graders; 52% male; and they were diverse racially and ethnically (42% Hispanic, 26% African-American, 20% White, and 11% another race/ethnicity category). A majority of these students (93%) had a positive view of the IYG lessons. Further, 83% of students said they were able to use the information and skills learned in IYG, and 86% of students were clearer about what they will/will not do regarding sex (data not shown).

See Appendix F (Table F.1) for more details on implementation quality.

3. Counterfactual experiences

Based on survey data from most comparison schools (7 schools in 7th grade and 8 schools in 8th grade out of 10 comparison schools), two schools reported providing sexual health education lessons in addition to the regular health, PE, or science curricula at the time comparison students in the study cohort were in 7th grade, and two schools provided sexual health education lessons when the students were in 8th grade. None of the comparison schools used an evidence-based curriculum during the two years of the study, although teachers noted they had used a set text or curriculum in the past, including *Worth the Wait*, *Big Decisions*, *Choosing the Best*; and the state-approved textbook. Other informal sexual health related activities included school-wide assemblies or presentations and a conference with a counselor/nurse (see Table F.2 in Appendix F).

4. Context

In terms of IYG implementation, no sites reported any substantial unplanned adaptations of the IYG curriculum.

Several unexpected changes in implementation plans occurred at four schools. At one intervention school, teachers informed the study team that they could not teach due to intense district oversight of a school improvement plan. To ensure implementation, staff facilitators from UT went to the school and taught the IYG curriculum. At three other intervention schools, the study team learned that a small number of the students did not receive the 7th grade curriculum due to schedule changes. To ensure complete implementation, staff facilitators from UT went to the schools and taught the 7th grade IYG curriculum to these identified 8th graders in the fall, prior to the regularly planned 8th grade curriculum implementation (see Table F.3 in Appendix F).

More details on other implementation evaluation findings can be found in Appendix F, including more detailed information on activities completed by lesson (Table F.4) and external events during the study time frame (Table F.5).

B. Impact study findings

Primary Behavioral Outcome: Initiation of Vaginal or Oral Sexual Intercourse. The rates of initiation of vaginal or oral sex at the end of 9th grade (final follow-up) estimated from the multilevel analysis model were 23.7% and 23.8% in the intervention and comparison conditions, respectively (Table IV.1); this difference was not statistically significant ($p = .956$). A sensitivity analysis examined the intervention effect without the covariate representing the differing school-wide rates of sexual initiation at baseline (Appendix G). The model-adjusted rates of initiation from this sensitivity analysis were 22.9% and 24.7% for intervention and comparison conditions, respectively, (Table G.1); consistent with the benchmark analyses, this difference did not reach statistical significance ($p = .37$).

Table IV.1. Post-intervention estimated effects using data from IYG student survey to address the primary research question

Outcome measure	Intervention %	Comparison %	Intervention compared to comparison mean difference (p -value of difference)
Initiation of vaginal or oral sex	23.7	23.8	-0.1 (0.956)
Sample Size ^a	804	739	

Source: Final follow-up survey, February to May of 9th grade, administered approximately 12 months after the program.

^aThe primary analytic sample was comprised of students who completed a baseline survey, a follow-up survey, provided values for covariates included in the final analysis models, and reported not having had vaginal or oral sex at baseline

Secondary Behavioral Outcome—Initiation of Vaginal Sex. The final adjusted rates were 19.7% (intervention) and 21.2% (comparison) at the 9th grade follow-up; this difference between conditions did not reach statistical significance ($p = .530$, see Table IV.2).

Secondary Behavioral Outcome—Oral sex at 9th grade. The final adjusted rates were 17.1% (intervention) and 14.7% (comparison) at the 9th grade follow-up; this difference between conditions did not reach statistical significance ($p = .241$, see Table IV.2).

Sensitivity analyses examined the intervention effects on the secondary outcomes without the covariate representing the differing school-wide rates of sexual initiation at baseline (Appendix E). The model-adjusted rates of initiation of vaginal intercourse from this sensitivity analysis were 19.0% and 22.1% for intervention and comparison conditions, respectively, (Table G.2); consistent with the benchmark analyses, this difference did not reach statistical significance ($p = .127$). The sensitivity analysis for the model assessing oral sex confirmed the benchmark, showing no statistically significant differences between conditions (see Table G.2 in Appendix G).

Table IV.2. Post-intervention estimated effects using data from IYG student survey to address the secondary research questions

Outcome measure	Intervention %	Comparison %	Intervention compared with comparison Mean difference (p -value of difference)
Initiation of vaginal sex	19.7	21.2	-1.5 (0.530)
Initiation of oral sex	17.1	14.7	2.4 (0.241)
Sample Size ^{a,b}	820, 829	762, 770	

Source: Final follow-up survey, February to May of 9th grade, administered approximately 12 months after the program.

^aThe first n represents the secondary analytic sample comprised of students who completed a baseline survey, a follow-up survey, provided values for covariates included in the final analysis models, and reported not having had vaginal sex at baseline.

^bThe second n represents the secondary analytic sample comprised of students who completed a baseline survey, a follow-up survey, provided values for covariates included in the final analysis models, and reported not having had oral sex at baseline.

V. Conclusion

This evaluation tested a replication of IYG in a similar geographic region to the original IYG studies. Data from the implementation evaluation suggest teachers delivered most lessons

and the majority of activities within each lesson. Further, among teachers reporting, most students received most lessons with good quality, but a large number of teachers did not provide attendance data despite incentives and extensive follow-up efforts, making it difficult to fully understand dosage across all students in the cohort. Further, there were a number of challenges throughout the study, such as teachers being pulled away from teaching IYG to focus on school improvement plans and students who had to make up 7th grade lessons at the beginning of the 8th grade year due to scheduling issues that may have contributed to gaps in exposure. Despite these challenges, both students and teachers reported positive reactions to the program, and most teachers expressed a desire to continue the program. In terms of behavioral impact, the results indicate there were no statistically significant differences in rates of vaginal or oral sexual initiation by the end of 9th grade.

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 initiation by type of sexual intercourse, including oral sex (Tortolero et al., 2010), anal (Tortolero et al., 2010) and vaginal intercourse (Markham et al., 2012).

A number of factors may have contributed to the pattern of behavioral results. First, the present study was an effectiveness trial using school teachers for curriculum implementation rather than outside facilitators as was done in the original IYG efficacy studies. The literature on efficacy and effectiveness research suggests that impact results are often diluted under effectiveness conditions, given the greater variation in implementation than would normally be found in efficacy trials (Glasgow, Lichtenstein, & Marcus, 2003). Second, the present study's primary outcome measure was initiation of vaginal or oral sex only rather than the combined

variable measuring initiation of oral, vaginal or anal sex used in the original IYG studies. The strongest effect in terms of reducing sexual initiation in the original IYG study was for anal sex, which was not measured in the present study because some study districts did not allow the question.

Future Analyses

This report represents the results of the primary and secondary behaviors only. Additional analyses have been completed and/or are underway to examine other critical intervention effects, such as on the theory-based psychosocial outcomes (e.g., knowledge, attitudes, and beliefs) as well as behavioral impacts on key subgroups, 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 has a number of strengths and limitations. Among its strengths, this study provides an effectiveness replication of an evidence-based program in a setting similar to the original research, thereby contributing to the literature on replication of evidence-based programs (EBP) using teachers rather than study staff. The study featured a randomized design involving 20 schools and long term follow-up of youth. All schools remained in the study throughout the length of the project, and the program was well received in the intervention schools.

Notable limitations include using a self-report questionnaire to collect outcome data, which is subject to potential response biases; nonetheless, some evidence supports the general reliability and validity of adolescents' reports of sexual and contraceptive behaviors, particularly

with the use of electronic devices (Trapl et al., 2005; Coyle et al., 2007; Palen, et al., 2008).

Additionally, implementation of the study's group-randomized trial design with only 20 units of randomization (schools) limits statistical power to detect significant differences, and resulted in some imbalances in school-wide rates of sexual behavior. These imbalances may not have been controlled fully with the school-level covariate used as a proxy to reflect the higher risk school environment presumed to exist at schools with higher overall rates of sexually active youth.

Missing data for IYG attendance for a notable percentage of sessions limits the conclusions on exposure for students in the treatment condition. Finally, the study included youth in urban middle schools in Texas; the results may not generalize to other geographic regions.

In conclusion, this study adds to the growing literature on replication of evidence-based programs, and underscores the need to better understand how varying aspects of the implementation affect the impact findings. Both the use of school teachers (as opposed to outside facilitators hired by the University of Texas Health Science Center at Houston and the change in how the outcome measure was operationalized could account for the differences in the pattern of findings. Further research on the role teachers play in the effectiveness of sex education programs and how to predict teacher needs for support prior to implementation may yield insights for continuing to strengthen training and support systems. It also may be beneficial to explore the impact of different implementation models in school settings, such as using a resource teacher (e.g., a health educator) or trained educators from community-based organizations. The results of this study also highlight the importance of systematically examining the impact of changing definitions in outcome measures when conducting replications; broader combination measures, such as including oral, vaginal, and anal intercourse, may be more sensitive with younger populations than relying on vaginal or oral sex alone.

VI. References

- Altman, D.G. (1991). *Practical Statistics for Medical Research*. Florida: Chapman and Hall/CRC Press.
- Centers for Disease Control and Prevention (CDC). *1991-2013 High School Youth Risk Behavior Survey Data*. Available at <http://nccd.cdc.gov/youthonline/>. Accessed on March 31, 2015.
- 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.
- Glasgow, R.E., Lichtenstein, E., & Marcus, A.C. (2003). Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. *Am J Public Health*, 93(8): 1261-1267.
- 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). 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. Available at http://tppevidencereview.aspe.hhs.gov/pdfs/Summary_of_findings.pdf. Accessed on February 1, 2015.
- Goldstein, H. (1995). *Multilevel Statistical Models*. London: Arnold Publishers.
- Graham, J.W., Flay, B.R., Johnson, C.A., Hansen W.B., & Collins, L.M. (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.
- Harris County Healthcare Alliance. *The State of Health in Houston/Harris County*. Available: <http://houstonstateofhealth.org/hshcms/wp-content/uploads/2013/04/State-of-Health-2013-Ex-Summary-Final.pdf>. Accessed on March 30, 2015.
- Hosmer, D.W. & Lemeshow, S. (1989). *Applied Logistic Regression*. New Jersey: John Wiley and Sons.
- 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.

Final: 7.14.16

Martin, J.A., Hamilton, B.E., Osterman, M.J.K, Curtin, S. C., & Mathews, M. S. et al(2015). Births: Final data for 2013. National Vital Statistics Reports, 64(1). Hyattsville, MD: National Center for Health Statistics. Available from:
http://www.cdc.gov/nchs/data/nvsr/nvsr64/nvsr64_01.pdf.

Martin, J.A., Hamilton, B.E., Ventura, S. J., Osterman, M.J.K., Curtin, S.C., & Mathews, T.J. (2013). Births: Final data for 2012. National Vital Statistics Reports, 62(9). Hyattsville, MD: National Center for Health Statistics. Available at
http://www.cdc.gov/nchs/data/nvsr/nvsr62/nvsr62_09.pdf

National Campaign to Prevent Teen and Unplanned Pregnancy. (2015). Texas Data.
<http://thenationalcampaign.org/data/state/texas>. Accessed on March 30, 2015.

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., Assmann, S.E., Enos, L.E., Kasten, L.E. (2002). Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. *Statistics in Medicine*, 21(19), 2917-2930.

Texas Department of State Health Services. (2012a). Birth Certificate Data File. Austin.

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.

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.

Ventura, S.J., Hamilton, B. E., & Mathews, T. J. (2014). National and state patterns of teen births in the United States, 1940-2013. *National Vital Statistics Reports*, 63 (4), 1-34.

Appendices

Appendix A: Data collection efforts

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

Data collection effort	Timing
Start date of programming	10/01/12-03/13
Baseline survey	09/24/12-3/6/2013
First follow-up (spring of 8 th grade; approximately 3 months post-program)	02/17/14-05/31/14 ^a
Final follow-up (spring of 9 th grade; approximately 12 months post-program)	02/05/15-05/31/15 ^b

^{a.} School-based surveying ended 5/31 each year; online, mail, and phone surveys continued through August.

^{b.} School-based surveying ended 5/31 each year; online, mail, and phone surveys continued through July.

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—Sessions delivered: How often were sessions offered? How many were offered?	Web-based implementation logs (assessed number of sessions, length of each session, date of sessions, among other elements)	Data were collected throughout implementation on all sessions. Teachers were expected to log sessions within 5 school days of teaching them, and were incentivized to log in a timely manner.	Evaluation staff; project staff monitor data for TA needs
Adherence—Dosage received: What and how much was received?	Attendance records	Student attendance was captured for each session in an Excel spreadsheet that was collected from IYG teachers at the end of the 12 sessions.	Evaluation staff; project staff monitor for completion
Adherence—Content covered: What content was delivered to youth?	Web-based implementation logs (specific activities completed, adaptations) Classroom observations	Data were collected throughout implementation on all sessions. Teachers were expected to log sessions within 5 school days of teaching them, and were incentivized to log in a timely manner. Classroom observation were completed on 3% of all sessions across facilitators (school years 2012-2013, 2013-2014, and 2014-2015)	Project staff; Evaluation staff
Adherence—Program staff and training: Who delivered material to youth?	List of facilitators from participating schools trained to implement program. List of facilitators from web-based implementation logs. Teacher survey	Training data were collected annually, with updates throughout school year if teachers transitioned. Web-based implementation logs were updated each semester. Teacher survey 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: Quality of staff-participant interactions	<p>Observations of interaction quality using required OAH observation protocol (Not a direct assessment of staff-participant interaction, but assesses rapport and communication with participants)</p> <p><i>On the following scale (1 = poor, 3 = average, 5 = excellent), rate the implementer on the following qualities...</i></p> <p><i>Comfort level discussing sex related topics e.g., reproductive anatomy, sex, condoms, contraception, teen pregnancy, STIs, etc.;</i></p> <p><i>Rapport and communication with participants; Effectively addressed questions/concerns</i></p>	<p>Sample of 3% of all sessions across 7th and 8th grade facilitators in 2012-2013, 2013-2014 and 2014-2015 school years.</p> <p><i>Note: We attempted to observe each teacher at least 2 times, teaching 2 different lessons and to spread observations equally across all lessons.</i></p>	Project staff; Evaluation staff
Quality: Quality of youth engagement with program	<p>Observations of engagement using OAH observation protocol</p> <p><i>How actively did the group members participate in discussions and activities?</i></p> <p><i>Scale: 1 = little participation, 3 = some participation, 5 = active participation</i></p>	<p>Sample of 3% of all sessions across 7th and 8th grade facilitators in 2012-2013, 2013-2014 and 2014-2015 school years.</p> <p><i>Note: We attempted to observe each teacher at least 2 times, teaching 2 different lessons and to spread observations equally across all lessons</i></p> <p>Student engagement was rated during each observation.</p>	Project staff; Evaluation staff
Counterfactual: Experiences of comparison condition	<p>Online health teacher survey</p> <p><i>Note: Survey focused on sexual health education (topics covered via class lessons and through other schoolwide events).</i></p>	<p>Annual survey of all teachers in counterfactual condition responsible for teaching health (7th grade completed in April 2013; 8th grade completed in April 2014).</p>	Evaluation staff
Context: Other TPP programming available or offered to study participants (both intervention and comparison)	<p>Meeting with District Coordinators.</p> <p>Teacher survey (online)</p> <p><i>Note: Survey focused primarily on other types of educational activities (e.g., assemblies). Could capture other TPP programming if teachers wrote in other TPP curricula.</i></p>	<p>District Coordinator meetings, twice per year.</p> <p>Annual survey of all teachers responsible for teaching health in intervention schools (May 2013, 2014, and 2015) and comparison conditions (2013 and 2014)</p>	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
Context: External events affecting implementation	News stories specific to study schools and districts. Record data (notes from internal meetings where issues were discussed)	Ongoing throughout the year (school and calendar year).	Project staff
Context: Substantial unplanned adaptation(s)	Web-based implementation log Record data (TA notes from meetings with sites, updated implementation plans showing substantial implementation changes); Observations of engagement using OAH observation protocol <i>Completion of Lesson Activities: Indicate which of the activities were completed during each class period by checking YES or NO. If an activity was not completed, please note which part of the activity was not completed and why.</i>	Log data are collected throughout implementation on all sessions. Teachers are expected to log sessions within 5 school days of teaching them, and are incentivized to log in a timely manner. Record data collected ongoing through the year. Sample of 3% of all sessions across 7th and 8th grade facilitators in 2012-2013, 2013-2014 and 2014-2015 school years. <i>Note: We attempted to observe each teacher at least 2 times, teaching 2 different lessons and to spread observations equally across all lessons. Student engagement was rated during each observation.</i>	Evaluation staff; project staff monitor data for adaptations

IYG = It's Your Game; OAH = Office of Adolescent Health; TA = Technical Assistances; TPP = Teen Pregnancy Prevention.

Appendix C: Study sample

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

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		20	10	10			
Clusters: Contributed at least one youth at baseline	<i>Baseline</i>	20	10	10	100	100	100
Clusters: Contributed at least one youth at follow-up	<i>3-months post-programming</i>	20	10	10	100	100	100
Clusters: Contributed at least one youth at follow-up	<i>12-months post-programming</i>	20	10	10	100	100	100
Youth: Eligible students in non-attributing clusters at time of baseline data collection ^a		3,565	1,818	1,747			
Youth: Who consented		2588	1,318	1,270	72.6	72.5	72.7
Youth: Contributed a baseline survey		2403	1,232	1,171	67.4	67.8	67.0
Youth: Contributed a follow-up survey	<i>3-months post-programming</i>	2,015	1,016	999	56.5	55.9	57.2
Youth: Contributed a follow-up survey	<i>12-months post-programming</i>	1,912	962	950	53.6	52.9	54.4

^a Clusters (i.e., schools) were randomly assigned to condition in fall 2011. Baseline data collection for the evaluation began in fall 2012. The sample includes youth who were enrolled in 7th grade at the school in the fall of 2012. These youth may have been at the school in fall 2011 at the time of random assignment (as 6th graders) or may have joined the school after fall 2011.

Appendix D: Equation for analytic sample

Equation for estimating benchmark analysis for primary outcome.

$$\text{Logit}(y_{jk(T3)}) = \beta_0 \text{cons} + \beta_1 \text{Tx}_k + \beta_2 \text{Evervagoral}_{k(T1)} + \beta_2 \text{District}_{k(T1)} + \beta_2 \text{SchoolEnr}_{k(T1)} + \beta_2 \text{FreeRed_lunch}_{k(T1)} + \beta_3 \text{Age}_{jk(T1)} + \beta_4 \text{Gender}_{jk(T1)} + \beta_5 \text{Black}_{jk(T1)} + \beta_6 \text{Latino}_{jk(T1)} + u_k + e_{jk}$$

where:

y_{jk} = 1 if student j in school k had vaginal or oral sex by final follow up; 0 otherwise

Tx_k = 0 if the school was in the control group, 1 if it was in the intervention group

$\text{Evervagoral}_{k(T1)}$ = school-level covariate representing proportion of students at baseline who had ever had vaginal or oral sex at school k

$\text{District}_{k(T1)}$ = school-level covariate representing district school k was in at baseline

$\text{SchoolEnr}_{k(T1)}$ = school-level covariate representing 7th grade enrollment in the year prior to the study start; design variable used for matched-pair randomization

$\text{FreeRed_lunch}_{k(T1)}$ = school-level covariate representing proportion of students at school who qualified for free/reduced lunch in the year prior to the study start; design variable used for matched-pair randomization

$\text{Age}_{jk(T1)}$ = age of student j in school k at baseline

$\text{Gender}_{jk(T1)}$ = 1 if student j in school k reported being female at baseline; 0 otherwise

$\text{Black}_{jk(T1)}$ = 1 if student j in school k reported being Black at baseline; 0 otherwise

$\text{Latino}_{jk(T1)}$ = 1 if student j in school k reported being Latino at baseline; 0 otherwise

β 's = unknown coefficients (to be estimated)

u_k = deviation of average for kth school's mean from overall mean

e_{jk} = deviation of average for jth student from kth school's mean

Appendix E: Implementation evaluation methods

Table D.1. Analysis methods used to address implementation research questions

Implementation element	Methods used to address each implementation element
Adherence—Sessions delivered: How often were sessions offered? How many were offered?	<p><i>The total number of sessions represents a sum of the sessions captured in the web-based project implementation log (Note: session = lesson).</i></p> <p><i>Average session duration was calculated as the average of the self-reported session lengths, measured in minutes (Note: session = lesson).</i></p> <p><i>Average weekly frequency was calculated as the total number of sessions divided by the total number of weeks when programming was offered (Note: session = lesson).</i></p> <p><i>(Note: A limitation of these data is that they are self-reported and subject to recall error. Some teachers did not turn in logs despite incentives and monitoring efforts, which may impact data quality).</i></p>
Adherence—Dosage received: What and how much was received?	<p><i>Average number of sessions attended was calculated as the average of the number of sessions that each student attended (Note: session = lesson).</i></p> <p><i>Average percentage of sessions attended was calculated as the total number of sessions attended divided by the total number of sessions offered (Note: session = lesson), averaged across all students in the intervention condition.</i></p> <p><i>(Note: A limitation of these data is that they are self-reported and subject to recall error. Some teachers did not turn in logs despite incentives and monitoring efforts, which may impact data quality).</i></p>
Adherence—Content covered: What content was delivered to youth?	<p><i>Data were collected on whether each lesson was delivered (curriculum includes 24 lessons), and the activities conducted in each lesson. Proportion of activities completed by lesson represents the total number of activities completed in a lesson with no or minor adaptations (per response categories on self-reported logs) divided by the total number of activities in that lesson (per the curriculum).</i></p> <p><i>(Note: A limitation of these data is that they are self-reported and subject to recall error. Some teachers did not turn in logs despite incentives and monitoring efforts, which may impact data quality).</i></p>
Adherence—Program staff and training: Who delivered material to youth?	<p><i>Total number of staff delivering the program represents a count of staff members implementing the program across schools each year of implementation. We also reported the positions of staff implementing the program using official school titles.</i></p> <p><i>% of staff trained was calculated as the # of staff members who participated in IYG trainings provided by UT staff divided by the total # of staff who delivered the program. Note: Training defined as: 7th Grade Level I IYG training (2 days) + 8th Grade Level II IYG training (1 day).</i></p> <p><i>(Note: A limitation of these data is that they are dependent on the quality of record keeping.)</i></p>

Implementation element	Methods used to address each implementation element
Quality: Quality of staff-participant interactions	<p><i>The data on quality of staff-participant reactions is presented as the number and percent of sessions coded for each response option on the five-point scale as well as a mean quality score. Further, an indicator of high-quality staff-participant interactions was calculated as the percent of sessions that were scored by the independent observer as a 4 or 5 on the rating scale (on a scale of 1-5 with 5 being most favorable) on the following indicators from the observation form--comfort level discussing sex related topics e.g., reproductive anatomy, sex, condoms, contraception, teen pregnancy, STIs, etc.; rapport and communication with participants; and effectively addressed questions/concerns.</i></p> <p><i>Note: To strive for a representative sample, we attempted to observe each computer lesson at least 2 times, and all other lessons at least 4 times. Additionally, each teacher was observed at least 2 times. Ultimately, however, the sample is one of convenience due to teachers' availability and thus, this measure may not be representative of all possible staff-participant interactions.</i></p>
Quality: Quality of youth engagement with program	<p><i>The data on quality of staff-participant reactions is presented as the number and percent of sessions coded for each response option on the five-point scale as well as a mean quality score. Further an indicator of the quality of youth engagement was calculated as the percent of sessions where the independent evaluator scored the following indicator as a 4 or 5: How actively did the group members participate in discussions and activities? Scale: 1, little participation, to 5, active participation.</i></p> <p><i>Note: To strive for a representative sample, we attempted to observe each computer lesson at least 2 times, and all other lessons at least 4 times. Additionally, each teacher was observed at least 2 times. Ultimately, however, the sample is one of convenience due to teachers' availability and thus, this measure may not be representative of all possible staff-participant interactions.</i></p>
Counterfactual: Experiences of counterfactual condition	<p><i>The data on the teacher survey assessing what health and sexuality education was taught at the schools in the counterfactual condition each school year is presented as frequency counts and percentages.</i></p> <p><i>(Note: A limitation of these data is that they are self-reported).</i></p>
Context: Other TPP programming available or offered to study participants (both intervention and counterfactual)	<p><i>All of the TPP programming available to both intervention and comparison groups described by district personnel or teachers via the teach survey.</i></p> <p><i>(Note: A limitation of these data is that they are self-reported).</i></p>
Context: External events affecting implementation	<p><i>The number of schools in which implementation was affected by district initiatives (unrelated to the TPP programming that occurred in this project).</i></p> <p><i>(Note: One limitation of these data is that they reflect issues brought to the attention of the project staff rather than a systematic assessment of external events that may impact implementation).</i></p>
Context: Substantial unplanned adaptation(s)	<p><i>The number of substantial unplanned adaptations and descriptions of the adaptations made.</i></p>

IYG = It's Your Game...Keep It Real; TPP = Teen Pregnancy Prevention.

Appendix F: Implementation findings

Table F.1. Analysis results of implementation adherence, quality, and context at intervention schools from teacher logs, teacher surveys, and classroom observation data

Implementation Element	7th Grade	8th Grade
	2012-2013	2013-2014
Adherence: Sessions Delivered -Number of sessions	10.4 out of 12 sessions 86.7% delivered across all classrooms ^a ^a Total n = 126.	11.7 out of 12 sessions 97.5% delivered across all classrooms ^a ^a Total n = 133.
Adherence: Sessions Delivered -average duration	45 minutes, range = 32-63	45 minutes, range = 31-110
Adherence: Sessions Delivered -average frequency of sessions	every 2.7 days range = 1.4-10.5	every 2.6 days range = .5-6.6
Adherence: Dosage Received -average number and % of sessions attended	Average = 10.9 or 90.8%	Average = 10.8 or 90.0%
Adherence: Dosage Received -percent of sample that did not attend any sessions	n = 0 ^b ^b Data missing for 662 participants.	n = 1, .1% of total sample without missing data ^b ^b Data missing for 508 participants.
Adherence: Content Covered average number and percentage of activities completed across all 12 lessons	Average number = 64, range 0-70 ^c Average = 95.3%, range = 0%–100% See Table F.4 for more detail. ^c Maximum total = 70.	Average number = 62, range 1-67 ^c Average = 93.0%, range = 11.1%-100% See Table F.4 for more detail. ^c Maximum total = 67.
Adherence: Program Staff & Training total number of staff delivering program	N = 27 teachers	N = 26 teachers
Adherence: Program Staff & Training average # of staff per school delivering program	Average = 3.0, range = 1-5 per school	Average = 2.6, range = 1-5 per school
Adherence: Program Staff & Training staff positions (official school titles)	70% PE teachers 30% other support staff or unknown	65% PE teachers 35% other support staff or unknown
Adherence: Program Staff & Training % of staff IYG trained	100%	100%
Adherence: Program Staff & Training % of staff receiving TA	100%	100%

Implementation Element	7th Grade	8th Grade
	2012-2013	2013-2014
Quality of Staff-Participant Interactions: Teacher comfort with topics Percent observed interactions with score of 4 or 5 out of 5 (1 = poor, 3 = average, 5 = excellent)	90.5% ^d ^d n = 52, 10.6% responded "n/a".	94.1% ^d ^d n = 47 observations. *27.7% responded "n/a".
Quality of Staff-Participant Interactions: Teacher rapport with students Percent observed interactions with score of 4 or 5 out of 5 (1 = poor, 3 = average, 5 = excellent)	86.5%	93.5%
Quality of Staff-Participant Interactions: Teacher ability to address questions Percent observed interactions with score of 4 or 5 out of 5 (1 = poor, 3 = average, 5 = excellent)	93.4% ^e ^e 6.3% responded "n/a" when there wasn't an opportunity for the teacher questions, e.g., questions weren't asked because it was a compute lesson.	96.5% ^e ^e 40.4% responded "n/a" when there wasn't an opportunity for the teacher questions, e.g., questions weren't asked because it was a compute lesson.
Quality of Youth Engagement with Program: Percent of sessions receiving score of 4 or 5 out of 5 on group participation (1 = little participation, 3 = some participation, 5 = active participation)	84.3%, n = 49 observations	72.4%, n = 47 observations

Table F.2 Analysis results of counterfactual at comparison schools from teacher survey

Counterfactual	7th Grade 2012-2013	8th Grade 2013-2014
Number of comparison schools providing sexual health education lessons in health, science, or PE classes	2 of 7	2 of 8
Number of comparison schools using an evidence based curriculum	0 of 7	0 of 8
Sexual health education topics addressed at schoolwide assemblies or events at comparison schools	# schools covering topic	# schools covering topic (n=8)
Healthy relationships	n/a	1
Decision making for health in general	n/a	2
Decision making for sexual health	n/a	1
Communicating values about sex	n/a	1
Identifying and avoiding risky sexual situations	n/a	1
Teen pregnancy	n/a	0
HIV/AIDS and other STIs	n/a	0
Abstinence	n/a	1
Condoms and/or contraception	n/a	0
Media influence on sexual health	n/a	1
Dating violence	n/a	1

n/a=not asked on the 7th grade teacher survey. These items were added to the teacher survey to assess this in 8th grade.

Table F.3. Analysis results of implementation context at intervention schools from teacher logs, teacher surveys, and classroom observation data

Implementation Element	7th Grade 2012-2013	8th Grade 2013-2014
Context: Substantial Unplanned Adaptations to Curriculum	0	0
Number of substantial unplanned adaptations		
Context: Substantial Unplanned Adaptations to Implementation	1 school	4 schools
UT facilitators taught lessons	See Table F.5 for more detail.	See Table F.5 for more detail.
Context: Other TPP Programming Available or Offered to Study Participants at Intervention Schools	At one school a guest speaker addressed certain information about sex.	None of the 25 teachers surveyed reported that sex education was offered at their sites through other (non-IYG) means.

Table F.4. Percentage of activities completed by lesson at intervention schools, 7th and 8th Grade based on teacher log data

7th Grade Lesson	N (logs)	Minimum	Maximum	Mean	Std. Deviation
1 (7 activities)	115	0	100	94.2	12.8
2 (7 activities)	113	85.7	100	99.0	3.7
3 (4 activities)	113	0	100	93.1	20.1
4 (7 activities)	108	71.4	100	97.5	6.1
5 (4 activities)	108	0	100	93.3	20.4
6 (8 activities)	108	87.5	100	98.7	3.8
7 (8 activities)	108	50.0	100	97.0	9.2
8 (3 activities)	108	0	100	95.4	17.9
9 (6 activities)	108	66.7	100	98.1	5.7
10 (5 activities)	108	0	100	92.0	20.6
11 (6 activities)	108	83.3	100	99.5	2.8
12 (5 activities)	108	80.0	100	96.3	7.8
Total Activities = 70				Average = 95.3%	

8th Grade Lesson	N (Logs)	Minimum	Maximum	Mean	Std. Deviation
1 (9 activities)	133	11.1	100	96.1	13.7
2 (6 activities)	130	50.0	100	97.7	7.1
3 (4 activities)	130	25.0	100	96.3	12.9
4 (6 activities)	129	50.0	100	95.9	9.3
5 (4 activities)	129	50.0	100	96.1	13.1
6 (6 activities)	129	50.0	100	96.8	9.6
7 (4 activities)	129	50.0	100	97.3	10.9
8 (7 activities)	129	42.9	100	94.5	13.5
9 (4 activities)	129	25.0	100	93.4	16.4
10 (9 activities)	129	44.4	100	94.7	12.7
11 (4 activities)	129	0	100	88.8	24.8
12 (4 activities)	129	0	100	90.7	22.5
Total Activities = 67				Average = 93.0%	

Table F.5. External events during study time frame from team meeting notes

Continuous Conditions

In September 2012, a small group of parents from a non-study district raised concerns about the planned adoption of the IYG curriculum in their schools. The concern was covered in the media and UT staff members and the developer were involved in many community meetings to answer questions and dispel myths about the nature of the curriculum. One protestor filed an open records request for all districts in Harris County related to sexual education. One protestor protested outside an IYG training. Despite this negative attention, all schools remained committed to their participation in the IYG study and schools in the comparison condition continued to request to teach the program after the study.

During the study cohort's 7th grade year

In one school, one class IYG was taught by UT staff because the PE teacher was not comfortable teaching IYG.

During the study cohort's 8th grade year

In one school, three classes IYG was taught by UT staff because teachers were asked to focus on program improvement status requirements.

In three schools, a group of 8th grade students did not receive the curriculum in 7th grade, so they received the 7th grade curricula prior to receiving the 8th grade curricula in 8th grade.

Severe weather (flooding) affected all areas schools and many schools were closed for several days, affecting planned implementation.

A fire in one school delayed planned implementation for several days.

ITYG = It's Your Game...Keep It Real

Appendix G: Sensitivity analyses

Table G1. Sensitivity of impact analyses using data from UT IYG to address the primary research question

Intervention compared with comparison	Benchmark approach difference	Benchmark approach <i>p</i> -value	Sensitivity approach excluding school-level % had sex covariate - difference	Sensitivity approach excluding school-level % had sex covariate <i>p</i> -value
Initiation of vaginal or oral sex	-0.1	0.956	-1.8	0.427
Sample Size	Intervention n: 804	Comparison n: 739		

Source: Final follow-up survey, spring of 9th grade, administered approximately 12 months post-program.

Notes: Sensitivity analysis excluded school-level percentage of students who ever had vaginal or oral sex covariate from the model. Difference represents difference between percentage of intervention and comparison I initiators.

Table G2. Sensitivity of impact analyses using data from UT IYG student survey to address the secondary research questions

Intervention compared with comparison	Benchmark approach difference	Benchmark approach <i>p</i> -value	Sensitivity approach excluding school-level % had sex covariate - difference	Sensitivity approach excluding school-level % had sex covariate - <i>p</i> -value
Initiation of vaginal sex	-1.5	0.530	-3.1	0.127
Initiation of oral sex	2.4	0.241	0.5	0.772
Sample Size ^{a,b}	Intervention n: 820, 829	Comparison n: 762, 770		

Source: Final follow-up survey, spring of 9th grade, administered approximately 12 months post-program.

Notes: Sensitivity analysis excluded school-level percentage of students who ever had vaginal or oral sex covariate from the model. Difference represents difference between percentage of intervention and comparison initiators.

^aThe first n represents the secondary analytic sample comprised of students who completed a baseline survey, a follow-up survey, provided values for covariates included in the final analysis models, and reported not having had vaginal sex at baseline.

^bThe second n represents the secondary analytic sample comprised of students who completed a baseline survey, a follow-up survey, provided values for covariates included in the final analysis models, and reported not having had oral sex at baseline.