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

Evaluation of *Seventeen Days* in Ohio, Pennsylvania, and West Virginia

Final Impact Report for

Carnegie Mellon University

August 31, 2015

Prepared by

Joan Eichner, MPA, MPH
Jennifer Salaway, PhD
Janell Smith-Jones, PhD
Robert McCall, PhD

University of Pittsburgh, Office of Child Development
400 North Lexington Street
Pittsburgh, PA 15208

Eichner, J., Salaway, J., Smith-Jones, J., & McCall, R. (2015). "Evaluation of *Seventeen Days* in Ohio, Pennsylvania, and West Virginia: Findings from the Replication of an Evidence-Based Teen Pregnancy Prevention Program." Pittsburgh, PA: University of Pittsburgh, Office of Child Development.

Acknowledgments:

We would like to thank the Carnegie Mellon University and West Virginia University teams for their partnership during this project, particularly Julie Downs, Pamela Murray, Vanessa Veltre, Mandy Lanyon, Amie Ashcraft, and Janie Leary. We would also like to thank Jean Knab at Mathematica for her guidance and support throughout the evaluation.

An additional thank you is given to Hoi Suen and his students, Wik Hung Pun and Alex Towle, from Penn State University, for their statistical knowledge and willingness to share their guidance and expertise to our team.

This publication was prepared under Grant Number TP1AH00040 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 *SEVENTEEN DAYS* IN OHIO, PENNSYLVANIA, AND WEST VIRGINIA: FINDINGS FROM THE REPLICATION OF AN EVIDENCE-BASED TEEN PREGNANCY PREVENTION PROGRAM

I. Introduction

A. Study overview

Despite the encouraging news that teen birth rates in the United States have declined substantially since the 1990s, with a decrease of 9% since 2013 alone (Hamilton, Martin, Osterman & Curtin, 2015), these rates remain higher in the U.S. than in other industrialized countries (Sedgh, Finer, Bankole, Eilers, & Singh, 2015). Approximately 750,000 teen pregnancies occur in the U.S. each year, of which 82% are unintended (Kost & Henshaw, 2012). Teen pregnancy and births generate enormous financial and society costs including medical care, public assistance, lost wages, and poorer educational outcomes for teen mothers, to name a few.

The U.S. Department of Health and Human Services (HHS) contracted Mathematica Policy Research to systematically review evaluations of programs targeting reductions in teen pregnancy, sexually transmitted infections (STIs), and risky sexual behavior in adolescents 19 and younger. Through its Teen Pregnancy Prevention Program, the Office of Adolescent Health at (HHS) invests in replications of effective evidence-based programs to address teen pregnancy. In 2010, Carnegie Mellon University in Pittsburgh, Pennsylvania received Tier 1 funding from the Office of Adolescent Health to update, replicate, and evaluate a video-based intervention called *What Could You Do?*, which had been determined to be of high quality during rigorous review (Goesling, Colman, Trenholm, Terzian & Moore, 2014). *What Could You Do?* was developed and evaluated by the Carnegie Mellon University research team and its partners in the 1990s (Downs, Murray, Bruine de Bruin, Penrose, Palmgren & Fischhoff, 2004). *What Could You Do?* met the Teen Pregnancy Prevention Evidence Review criteria for a high-quality study

with short-term impact, defined as positive effects for the entire sample and effects lasting less than a year. This rating indicates that the original study used random assignment, met acceptable standards for sample attrition, controlled for differences between groups at the start of the study, and used the same data collection methods for intervention and comparison groups (Mathematica Policy Research & Child Trends, 2010). *What Could You Do?* was found to have a statistically significant program impact on reducing sexual activity, condom failure, and sexually transmitted infections, but not on attempted contraceptive use. The evaluation did not measure pregnancy outcomes (Goesling, Colman, Trenholm, Terzian & Moore, 2014).

With Tier 1 funding from the Office of Adolescent Health, *What Could You Do?* was re-created in 2011 as *Seventeen Days,* using updated technology and minimal modifications, such as an enhanced production quality, increased racial and ethnic diversity of the cast, updated medical content, expanded coverage of birth control options, and an emphasis on the risk of unplanned pregnancy. Both interventions have similar critical features: they are interactive videos delivered individually that provide cognitive rehearsal opportunities for viewers to practice decision-making in sexual situations. *Seventeen Days* aims to reduce risky behaviors associated with sexually transmitted infections and unintended pregnancy, increase knowledge of risk, and increase adolescents' knowledge of sexual health and safe sexual behavior.

To test the effects of *Seventeen Days*, the research team conducted an evaluation in a tristate area including Western Pennsylvania, Franklin County, Ohio, and multiple locations throughout West Virginia. These areas were selected due to their relatively high teen birth rates and to provide a mix of urban and rural populations.

B. Primary research questions

The evaluation sought to test the effectiveness of *Seventeen Days* in reducing risky sexual behavior in adolescent females. The primary research question was:

Does the *Seventeen Days* interactive video result in safer sexual behavior (defined as abstinence or no reports of vaginal sex events unprotected by a condom) for adolescent females at six months post-randomization relative to the control video?

The primary behavioral outcome of interest, safer sexual behavior, was operationalized as an adolescent being abstinent or reporting no vaginal sex events unprotected by a condom during the three-month period prior to the data collection that took place six months after randomization.

C. Secondary research question(s)

Additionally, five secondary research questions allowed for a more nuanced measure of the impacts of *Seventeen Days.* The secondary research questions were:

1. Does the *Seventeen Days* interactive video result in safer sexual behavior (abstinence or no reports of vaginal sex events unprotected by a condom) for adolescent females at 3 months after randomization relative to the control video?
2. Does the *Seventeen Days* interactive video result in abstinence from sexual behavior among female adolescents at 3 months after randomization relative to the control video?
3. Does the *Seventeen Days* interactive video result in abstinence from sexual behavior among female adolescents at 6 months after randomization relative to the control video?
4. Does the *Seventeen Days* interactive video result in a lower proportion of positive pregnancy test results reported at 6 months after randomization relative to the control video?
5. Does the *Seventeen Days* interactive video result in a lower proportion of positive infection (Chlamydia and Gonorrhea) test results at 6 months after randomization relative to the control video?

As with the primary research questions, the measures for these questions refer to behaviors and experiences during the three-month period prior to the data collection.

II. Program and comparison programming

A. Description of program as intended

The *Seventeen Days* intervention, which was adapted from the evidence-based *What Could You Do?* intervention, is a stand-alone interactive video that can be used in health clinics or other settings to promote safer sexual behaviors and decision-making among adolescent females. The overarching goals of the program are to reduce both teen birth rates and rates of sexually

transmitted infections among teenage girls. These goals will be achieved through a series of intermediate steps starting with the implementation of the activities in the intervention (character vignettes, gynecological exam, condom demonstration, disease information, birth control information), and potentially mediated by both individual (e.g. age, relationship status) and site-specific implementation factors (e.g. staff support and encouragement to participate). In the short-term, the intervention aims to increase self-efficacy around abstinence and condom use, and increase error-free, consistent condom use, and abstinence from sexual behaviors. These intermediary changes lead to fewer positive pregnancy and Chlamydia and Gonorrhea test results. The intervention logic model is depicted in Appendix A.

The video is individually self-administered by girls via an electronic tablet in the clinic while they are waiting for care, and it can be continued via any Internet-enabled device with a Flash player outside of the clinic.[1] Therefore, girls may start at the clinic and finish elsewhere, or even complete all their entire viewing outside of the clinic. The viewing experience is individualized, because participants have their own tablet and can select content relevant to their needs without being inhibited or influenced by other people. The video also invites the girls to apply the demonstrated skills in their own lives.

The video features several interactive components. Viewers are able to choose the type of material they wish to view from a menu of options, including the selection of one or more characters whose narrative they can view. These narratives offer decision points where viewers choose one of three options and then watch the character enact the chosen option. Finally, each

[1] In the clinic, recruiters told girls how they could log in and watch the video later, and the deadlines to complete viewing. If girls expressed Internet or computer access challenges, recruiters (or members of the study team by text or email) helped girls by looking up libraries in their area where they might access the computer or by suggesting possible businesses with free WiFi that may be accessible (e.g., McDonald's).

narrative story offers "cognitive rehearsal" points where the action pauses and the viewer has time to mentally rehearse how she would respond in the situation depicted by the character.

Seventeen Days consists of about 2.5 total hours of video material. The core material of *Seventeen Days* lasts approximately 35 minutes and consists of one character (Jessica) narrating an introduction to the concepts (of choice in sexual situations and cognitive rehearsal of safe choices), a lesson on condom procedure and efficacy, and a vignette on sexual negotiation, including choices and cognitive rehearsal (Table 2.1). Viewers are also provided the opportunity to choose any of the other six character's stories. After completing the core content, participants have the option to return to the video materials as often as they like over the course of six months after enrolling in the study. Viewers are not able to choose additional material to view until they have completed the core viewing material and can go back to watch any portion of the video again, via the Return Menu.

Table 2.1. Core features and additional material in *Seventeen Days*

Features of Seventeen Days	Core Viewing Material	Additional Viewing Material
Jessica's Story: opening scenes, vignette with decision-making points, and conclusion.	X	
A condom demonstration scene providing background on both how and why condoms lower risk	X	
A set of vignettes depicting other characters modeling skills for negotiating lower-risk sexual behavior (including abstinence) with different kinds of partners		X
Mini-documentaries including: a gynecological exam explaining female physiology and modeling interactions with a health care provider, including how to ask for further services; explanation of the difference between viral and bacterial infections, the difficulty of identifying infections in partners or oneself, health consequences, and treatment options, if any; a brief anatomy lesson identifying the internal and external female genitalia; and descriptions of various birth control options, how they work and why they are important to use		X

B. Description of counterfactual condition

Like the intervention video, *Driving Skills for Life*, the comparison condition, is a stand-alone interactive video individually self-administered by adolescent girls via an electronic tablet. Produced by the Ford Motor Company for adolescent drivers, the video provides information on how to reduce driving risks. As with *Seventeen Days*, participants can watch the video using a tablet while they are waiting for care during a clinic appointment and continue to access the video on any other Internet-enabled device at their own pace. Girls can select content relevant to their needs without being inhibited or influenced by other people.

The core material lasts approximately 35 minutes; girls may watch it in one or more viewing sessions. It includes introductory lessons in car handling such as how to handle different road conditions, followed by an interactive menu in which the viewer can choose among different types of content, including short instructional videos and interactive games (e.g., practice merging onto a busy highway by managing speed with keyboard keys and monitoring other cars on screen). After this core material, girls have the option to return to the video materials as often as they like over the course of six months of unlimited access. There is about 2.5 hours of total driving material. The length of core programming and amount of total amount of material are approximately equal to *Seventeen Days.*

III. Study design

An individual randomized controlled trial was used to assess the impacts of the *Seventeen Days* interactive video intervention on reducing risky sexual behaviors among adolescent females in three Mid-Atlantic and Midwestern states. The design allows causal attribution of impacts of the intervention itself because random assignment between conditions ensures any differences are the result of random chance, not systematic differences between the groups.

A. Sample recruitment

Adolescent females were recruited from 20 participating clinics in Columbus, Ohio, Western Pennsylvania, and throughout West Virginia, from June, 2012 to December, 2014. These clinics represented a range of types of medical care facilities and geographic areas. One site was an urban hospital specializing in women's health, three sites were hospital-based adolescent medicine clinics, six sites were family planning clinics, and ten sites were county health department family planning clinics.

The study design aimed to test the intervention as part of normal operations in clinics that are frequently used by adolescents. This setting was similar to that of the original evaluation of *What Could You Do?* **(**Downs, Murray, Bruine de Bruin, Penrose, Palmgren & Fischhoff, 2004), and included two of the same clinics in this evaluation. Participating clinics in the tristate area were identified by the principal investigators, who visited each clinic site early in the first grant year to build relationships and develop an implementation plan tailored to each site that matched the needs and context of each clinic's staff, client population, physical space, clinic schedule, and resources.

During the recruitment phase, paid research staff recruited regularly in the highest volume clinics and the project team visited smaller clinics on days when potentially eligible patients were scheduled or anticipated. Recruiters approached any young woman in each clinic's waiting room who appeared to be within the target age range. They described the study as the opportunity to watch videos that can help teens make choices about driving and about relationships. If the adolescent was interested, she was given a study tablet with a short screening survey. Eligibility requirements included being female, aged 14–19, sexually active in the past six months, and not currently pregnant. If the teen was eligible, the system directed her to more information and the consent and enrollment process. All Institutional Review Boards granted a

waiver of parental consent for minors who were in the clinic without a parent or guardian accompanying them inside the building. If the parent was present, parental consent was obtained for minors and minors provided assent. Adolescents ages 18 and 19 consented for themselves.

Of the 5,272 girls screened, 53% were eligible to participate in the study (n=2,814). Of those eligible, 70% consented to participate (n=1,957). The final sample consisted of 1,317 adolescent girls who were patients at one of the 20 participating clinics and were randomized into intervention or comparison conditions.

B. Study design

The design was an individual randomized controlled trial in which each participant had a 50/50 probability of assignment to the intervention (*Seventeen Days*) or comparison (*Driving Skills for Life*) video group. Immediately after consent and the completion of the baseline survey measures, participants were randomly assigned to a condition by a computer program and automatically routed to the appropriate video via an electronic tablet. To ensure that assignment to both conditions occurred similarly, randomization occurred at each clinic.

C. Data collection

1. Impact evaluation

The sources of data for the impact evaluation include an electronic survey and clinical test results. Behavioral data were obtained electronically through a survey that was completed at baseline, three months post-randomization, and six months post-randomization. Electronic surveys had skip-logic that branched questions based on unique responses and allowed a large amount of data to be filled automatically based on responses.

Once consent had been given, each adolescent created a username and password, and the tablet took her to the baseline survey. The survey began with a detailed self-report measure of sexual behavior using the format of a Timeline Followback calendar, which asked each

participant to list the number of times she had sex in the preceding three months, if this was a new or regular partner, the type of sexual activity (vaginal, oral, anal), and the number of those times a condom was used with or without problems. The calendar tool has been found to produce fairly accurate reporting by adolescents regarding their sexual behavior up to three months prior, with high test-retest reliability (Cronbach alpha .86 to .97) (Weinhardt et al., 1998) (Sobell & Sobell, 1992).

Once the calendar portion of the online survey was finished, the participant was directed automatically to a test of knowledge about condom use and sexual health. This section was followed by a series of self-efficacy questions regarding different aspects of sexual behavior, such as negotiating safer behaviors with a partner, deciding against sex with a partner, and using condoms correctly and consistently. Last, the survey included questions regarding the incidence of pregnancies and sexually transmitted infections (STIs). Demographic data were also collected in this survey. The three- and six-month surveys collected the same information except for demographics.

Following the baseline survey and automatic randomization, participants were directed to the appropriate video, which was queued to start with the core video material. The teen could carry the tablet as she moved from waiting room to exam room, and the session could be paused and resumed as needed at the clinic appointment or later from a personal device or on a study tablet at the clinic.

At three- and six-months post-randomization, participants completed the follow-up online surveys. Project staff called, emailed, texted, and/or mailed reminders to participants when their follow-up surveys were due to be completed. Participants were instructed to complete them on

their own in any location with Internet access, or they could return to complete them at the clinic using study tablets. All data collection procedures were identical in both conditions.

Clinical outcome data on pregnancy and STIs were collected via self-administered vaginal swabs at baseline and six months post-randomization. Study recruiters provided participants with a kit to test for baseline clinical infection with Chlamydia or Gonorrhea and a home pregnancy test. Participants were instructed to complete these in the clinic on the day of enrollment. If unable or unwilling to collect in the clinic, they were permitted to collect at home. If completed at home, they were provided with return packaging and postage so that the kits could be returned to a central location for processing. Follow up clinical testing (at six months post-randomization) could also be conducted at the clinic or completed at home using a kit mailed by study recruiters.

Participants were given $25 gift cards for completing the surveys and clinical test kits during the appropriate times. At the six-month point, they received a $50 gift card for completing all data collection and a $20 bonus if their previous data were submitted on time.

2. Implementation evaluation

Because the intervention is a self-administered online video, it is delivered with fidelity at each viewing, therefore study implementation was measured in terms of participants' viewing choices. Understanding this viewing behavior is important because the amount of viewing time is essentially the "dose" of the intervention, and viewing specific content or more content overall is hypothesized to be associated with participants' outcomes. Those who chose to view content multiple times increased their overall dosage. Unfortunately, the viewing data do not allow a measure of what material was viewed more than once by each participant.

The viewing data for both the intervention and comparison groups provide the number of views each participant made during her six-months of access to the video, whether the core viewing was completed, the viewing choices each participant made, and what if any additional

content (i.e., content beyond core material) was viewed. All viewing data were collected automatically in real time via the Internet-enabled device used by participants to watch the video. Data were also collected on external events affecting implementation such as technology issues and time to complete the program. Information on alternative sources of sexual education, such as schools, was not collected. Participants attended junior and senior high schools in many school districts within the catchment areas of the 20 clinics involved. Schools' sexual health curricula vary by grade, district, and state; therefore, the nature and amount of alternative sexual education received among participants would have been highly variable. See Table B.1 in Appendix B for a description of the data used to address implementation questions.

D. Outcomes for impact analyses

The behavioral outcome of interest for the primary impact analysis research question, safer sexual behavior, was determined by either 1) no self-reported sexual activity within the preceding three months, or 2) no self-reported vaginal sexual events unprotected by a condom within the preceding three months (Table 3.1). These data were reported by participants on the electronic Timeline Followback calendar tool in the survey administered at six-month post-randomization. A dichotomous outcome measure shows which teens engaged in safe behavior either through abstinence or by using condoms during sexual activity.

The secondary research questions were also answered using behavioral outcomes data collected via the Timeline Followback calendar as well as clinical tests. For these questions, safe sexual behavior was determined by either 1) no self-reported sexual activity within the preceding three months, or 2) no self-reported vaginal sexual events unprotected by a condom within the preceding three months (Table 3.2). Pregnancy and STI rates were obtained through self-administered test kits at six months post-randomization, as shown in Table 3.2.

Table 3.1 Behavioral outcomes used for primary impact analyses research question

Outcome name	Description of outcome	Timing of measure relative to program
Safe sexual behavior	This variable is a yes/no measure of whether a participant engaged in safe sexual behavior within the preceding 3 months. The measure is constructed from the following 2 items on the calendar: • Have you had vaginal sex in the last 3 months? • What protection did you use? Respondents who respond "no" they have not had vaginal sex or respondents who respond "yes" they have had vaginal sex *and* no reports of vaginal sex events unprotected by a condom are coded as 1 and all other respondents are coded as 0.	6 month post-randomization

Table 3.2. Behavioral outcomes used for secondary impact analyses research questions

Outcome name	Description of outcome	Timing of measure relative to program
Safe sexual behavior	This variable is a yes/no measure of whether a participant engaged in safe sexual behavior within the preceding 3 months. The measure is constructed from the following 2 items on the calendar: • Have you had vaginal sex in the last 3 months? • What protection did you use? Respondents who respond "no" they have not had vaginal sex or 2) respondents who respond "yes" they have had vaginal sex *and* no reports of vaginal sex events unprotected by a condom are coded as 1 and all other respondents are coded as 0.	3 months post-randomization
Abstinence from sexual behavior	This variable is a yes/no measure of whether a participant abstained from vaginal sexual behavior within the preceding 3 months. The measure is taken directly from the following item on the calendar: • Have you had vaginal sex in the last 3 months? Respondents who respond "no" are coded as 1 and respondents who respond "yes" are coded as 0.	6 months post-randomization
Reported pregnancies	This outcome is the result of the pregnancy test (positive or negative) in the test kit data. The variable is constructed as a dichotomous variable in which a positive test result is coded as a 1 and a negative test results is coded as a 0.	6 months post-randomization

Outcome name	Description of outcome	Timing of measure relative to program
Infection test results	This outcome is the result of the clinic test for Chlamydia and Gonorrhea (positive or negative) in the test kit data. The variable is constructed as a dichotomous variable in which a positive test result for Chlamydia or a positive test result for Gonorrhea is coded as a 1 and a negative test result is coded as a 0.	6 months post-randomization

E. Study sample

For the primary research question, the final analytic sample included all participants who were randomized to a condition and completed a six-month online survey. For the secondary research questions, the analytic sample included all participants who were randomized and completed either the three-month or six-month survey, or test kit, depending on the question. The final analytic sample for the six-month survey was 674 participants (51% total response rate). The final analytic sample for the three-month survey was 685 participants (52% response rate), and 563 participants for the test kit at six months (43% response rate). There was no differential attrition between the intervention groups. See Table C.1 in Appendix C for sample sizes by intervention status.

F. Baseline equivalence

The evaluation team conducted baseline equivalence tests for the three- and six-month analytic samples to assess whether attrition affected the comparability of the intervention and comparison groups and to examine the differences between the groups on the following demographic variables and behavioral measures: age, race, ethnicity, pregnancy history, STI history, and baseline measures of abstinence and safe sexual behavior for each analytic sample. The analytic approach used a linear probability model. Tables 3.3 to 3.5 summarize the key baseline measures for the three analytic samples. There are no significant differences ($p < .05$)

between the intervention and comparison groups on the key baseline characteristics for any of the analytic samples.

Table 3.3. Summary statistics of key baseline measures for youth completing 6-month Timeline Followback Calendar survey (sample for primary outcome analysis)

Baseline measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison *p*-value of difference
Age	17.267 (1.492)	17.215 (1.457)	0.052	0.649
Gender (female)	1.000	1.000	0.000	1.000
Race: White	0.567	0.584	0.017	0.966
Race: Black	0.329	0.310	0.019	0.783
Race: Asian*	0.003	0.00	0.003	--
Race: American Indian/Alaskan Native	0.003	0.006	0.003	0.596
Race: More than one race	0.098	0.099	0.001	0.735
Ethnicity: Hispanic	0.057	0.050	0.007	0.705
Ever been pregnant	0.099	0.128	0.029	0.240
Ever been infected with Chlamydia or Gonorrhea	0.098	0.058	0.040	0.058
Abstinent in past 3 months	0.097	0.098	0.001	0.969
Safe sexual behavior in past 3 months	0.306	0.306	0.000	0.990
Sample size	334	340		

Notes. Analytic sample size reflects those with non-missing values on the primary outcome measure. *N=1 for this category; no statistical test conducted.

Table 3.4 Summary statistics of key baseline measures for youth completing 3-month Timeline Followback Calendar survey (sample for secondary outcome analysis)

Baseline measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison *p*-value of difference
Age	17.231 (1.492)	17.295 (1.444)	0.064	0.572
Gender (female)	1.000	1.000	0.000	1.000
Race: White	0.576	0.589	0.013	0.850

Baseline measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison *p*-value of difference
Race: Black	0.321	0.297	0.024	0.602
Race: Asian	0.003	0.003	0.000	0.966
Race: American Indian/Alaskan Native	0.006	0.009	0.003	0.715
Race: More than one race	0.095	0.102	0.007	0.959
Ethnicity: Hispanic	0.050	0.056	0.006	0.719
Ever been pregnant	0.120	0.121	0.001	0.966
Ever been infected with Chlamydia or Gonorrhea	0.070	0.061	0.009	0.615
Abstinence	0.081	0.091	0.010	0.664
Safe Sexual Behavior	0.292	0.316	0.024	0.850
Sample size	342	343		

Note. Analytic sample size reflects those with non-missing values on the secondary outcome measure.

Table 3.5. Summary statistics of key baseline measures for youth completing 6-month Clinic Test Kit (sample for secondary outcome analysis)

Baseline measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison *p*-value of difference
Age	17.311 (1.496)	17.251 (1.455)	0.060	0.631
Gender (female)	1.000	1.000	0.000	1.000
Race: White	0.565	0.535	0.030	0.216
Race: Black	0.351	0.360	0.009	0.345
Race: Asian*	0.004	0.000	0.004	--
Race: American Indian/Alaskan Native*	0.004	0.000	0.004	--
Race: More than one race	0.076	0.105	0.029	0.361
Ethnicity: Hispanic	0.061	0.054	0.007	0.709
Ever been pregnant	0.107	0.118	0.011	0.689
Ever been infected with Chlamydia or Gonorrhea	0.097	0.065	0.032	0.174

Baseline measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison *p*-value of difference
Abstinence	.0.117	0.105	0.012	0.654
Safe Sexual Behavior	.0.320	0.320	0.000	0.910
Sample size	280	283		

Note. Analytic sample size reflects those with non-missing values on the secondary outcome measure. *N=1 for this category; no statistical test conducted.

G. Methods

1. Impact evaluation

We used an intent-to-treat framework and data collected at six months post-randomization to estimate the impact of *Seventeen Days* compared to the comparison group on participants' engagement in safe sexual behavior. An intent-to-treat analysis estimates the impact of the intervention on all participants who were randomized to a condition, regardless of whether they completed all or any of the *Seventeen Days* video content.

The analytic approach used a linear probability model to compare safe sexual behavior between the intervention group and the control group. The impact estimate is the regression-adjusted difference between the outcomes in both conditions. Impact estimates with *p*-values less than .05 (two-tailed tests) are considered statistically significant and provide evidence of a true difference between the groups as a result of *Seventeen Days*. Baseline covariates included demographic variables (age, race, and ethnicity) and the baseline value of the outcome variables. See Appendix D for model specifications.

Missing data occurred at both baseline and follow-up data collection points. To account for missing baseline covariates, we applied the dummy variable method. For our primary analysis of safe sexual behavior, when participants reported having sex in the last three months and then did not provide a follow-up protection response, we made the assumption that they did

not use a condom. For the three-month sample, the assumption was made for 154 out of 685 cases (77 in intervention and 73 in comparison); for the six-month sample, the assumption was made for 150 out of 674 cases (84 in intervention and 70 in comparison).

The analytic approach for the secondary research questions is the same approach applied to answering the primary research question with the exception of how we handled missing data. For our secondary outcomes (abstinence and clinic test results), our analytic samples only included participants who answered the abstinence question on the calendar and those participants who completed a test kit, respectively.

To determine if the results were sensitive to the analysis approach, we conducted additional analyses using alternative approaches (Appendix E). These included (1) analyses to compare the differences between the two groups without controlling for baseline covariates, (2) analyses that only included participants who provided responses to protection questions, and (3) analyses that treated missing protection responses as missing by creating a dummy variable for the "no protection used" assumed responses.

2. Implementation evaluation

Implementation of a video intervention is unique because the full video content was offered to every participant and the intervention is delivered with fidelity to every participant; however, dose varies because each participant chooses content and number of viewings. For both conditions, data on viewing behavior was collected electronically in real time to assess dosage.

Frequency analyses by intervention group were run on the viewing data to determine what material was viewed. Additional description of the data and its operationalization used in the implementation analysis can be found in Appendix E Table E.1.

IV. Study findings

The impact evaluation sought to determine if being offered the chance to view *Seventeen Days* resulted in safer sexual behavior, either abstinence or condom use, for adolescent females, relative to being offered the chance to view *Driving Skills for Life*, at two time points after randomization. Although the full dosage of intervention was offered to everyone, not all participants viewed the same amount of material or content. To understand the effects of the intervention, the implementation study examines what and how much video content was viewed.

A. Implementation study findings

Of the study participants in the intervention (*Seventeen Days*) group, 61% completed the entire core material that was determined to be necessary to have "seen the intervention". Of participants in the comparison group, 73% completed the core viewing of driving material. Percentages of core material viewed by participants in each group are summarized in Table 4.1.

Table 4.1. Percent of core material viewed by group.

Seventeen Days group			Comparison group		
Percent of core material viewed	**Percent of randomized participants**	**Percent of analytic sample**	**Percent of core material viewed**	**Percent of randomized participants**	**Percent of analytic sample**
0%	14.5%	0.3%	0%	19.0%	1.2%
10%	4.3%	0.3%	25%	2.7%	0.3%
50%	11.8%	9.6%	50%	3.2%	0.6%
75%	1.1%	0.6%	75%	2.3%	0.0%
90%	7.5%	7.2%	100%	72.9%	97.9%
100%	60.8%	82.0%			
Sample Size	653	334		664	340

Source. Data were derived from real time viewing records.

Note. For the different interventions, the content was broken up into different segments. For the comparison group, the control video was divided into four equal segments. For *Seventeen Days*, the natural cut-points did not directly match the four equal segments, so different cut-points were used to determine dosage.

Participants had the option to watch additional video, and more than one-half of each group did. In the *Seventeen Days* group, 55% watched additional video beyond core material. In

the comparison group, 52% watched some additional content. Over half of the participants viewed additional *Seventeen Days* vignettes, and 41% viewed additional mini-documentaries with cognitive rehearsal of negotiating risky sexual situations. The median number of vignettes watched was two. Twenty percent chose to view additional mini-documentary footage. The most frequently viewed (n=60) was the anatomy lesson while the least frequently viewed (n=0) was the birth control options. The most frequently viewed optional driving video material covered "space management" and "speed/management and safety tips".

Several events impacted the implementation. When 20 participants logged in a second time a programming glitch directed them to view the incorrect condition's video (that is, someone in the intervention group was provided a link to the comparison group video, and vice-versa). Only three of these participants were actually included in the analytic sample, however, because the other 17 did not complete the six-month survey. Additionally, every clinic in the study had shorter than anticipated wait times so most participants were only able to complete study registration before their appointment started. That meant participants had to stay after their appointment to finish watching the core material or log in again later. Finally, technology issues affected implementation at every clinic in the study. These issues varied by site but the major issues were slow and unreliable Internet connections in the rural sites, and conflicts between the clinic's and the study's wireless devices in some of the larger sites. Recruiters began to use MiFi devices purchased by the research team to achieve better connections.

B. Impact study findings

Table 4.2 shows the estimated effect of *Seventeen Days* on the primary outcome measure. We found no evidence that viewing *Seventeen Days* impacted engaging in safe sexual behavior compared to the comparison group. At the six-month post-randomization assessment, 37% of

intervention group members reported engaging in safe sexual behavior in the past three months, compared to 38% of the comparison group (p = .837).

Table 4.2. Post-intervention estimated effects using data from 6 month Timeline Followback Calendar survey to address the primary research question

Outcome measure	Intervention %	Comparison %	Intervention compared to comparison % difference (p-value of difference)
Safe sexual behavior	.371	.382	0.007 (0.837)
Sample Size	334	340	

Source: Timeline Followback Calendar administered 6 months post-randomization.
Notes: Safe sexual behavior is defined as either abstinent from sexual activity or no reports of vaginal sexual events unprotected by a condom in the last 3 months.

Secondary research questions

Tables 4.3 and 4.4 summarize the findings for the secondary research questions. First, we found no evidence that *Seventeen Days* impacted safe sexual behavior compared to the comparison group at the three-month post-randomization assessment. While this finding indicates a 2.5 percentage point difference on this outcome favoring the intervention group, this difference was not statistically significant (p = .402*)*, and as found in the primary outcome analysis, the difference marginally favored the control group at the six-month post-randomization assessment. Second, there is no statistically significant difference between the percentage of intervention (19.9%) and comparison (20.4%) group members abstaining from sexual activity at three months post-randomization or between the percentage of intervention (21.5%) and comparison (20.1%) group members abstaining from sexual activity at six months post-randomization. Finally, we did not detect an impact for *Seventeen Days* on either pregnancy or infection rates at the six-month post-randomization assessment. While there was less than a half percentage point difference on pregnancy favoring the intervention group; the difference was not statistically significant. There is no statistically significant difference in the percentage

of intervention (8.6%) and control (7.4%) group members having a positive infection test at six months post-randomization.

Table 4.3. Post-intervention estimated effects using data from Timeline Followback Calendar survey to address the secondary research questions

Outcome measure	Intervention %	Comparison %	Intervention compared with comparison % difference (*p*-value of difference)
Safe sexual behavior at 3 months	0.383	0.364	0.029 (0.402)
Abstinence at 3 months	0.199	0.204	0.001 (0.982)
Abstinence at 6 months	0.215	0.201	0.008 (0.790)
Sample Size (3 months)	342	343	
Sample Size (6 months)	334	340	

Source: Timeline Followback Calendar administered in the survey at 3 or 6 months post-randomization.
Notes: Safe sexual behavior is defined as either abstinent from sexual activity or no reports of vaginal sexual events unprotected by a condom in the last 3 months.

Table 4.4. Post-intervention estimated effects using data from Clinic Test Kit to address the secondary research questions

Outcome measure	Intervention %	Comparison %	Intervention compared with comparison % difference (*p*-value of difference)
Positive pregnancy test result	.014	.018	0.004 (0.723)
Positive infection test result	.086	.074	0.005 (0.817)
Sample Size	280	283	

Source: Clinic Test Kit collected at 6 months post-randomization.
Notes: Clinic Test Kits tested for Chlamydia and Gonorrhea.

Across all sensitivity analyses, findings were consistent with those found using the benchmark approach (Appendix E Tables E.1 and E.2). In addition, the primary analysis was conducted again excluding participants (n = 3) who had viewed the video of the alternate condition and provided a six-month post-randomization survey; findings were consistent with

those found for the entire analytic sample. Future exploratory analyses will examine the outcome of only those participants who completed the core viewing in each condition.

V. Conclusion

This evaluation is the first rigorous study of the *Seventeen Days* interactive video, an updated and expanded adaptation of *What Could you Do?,* a video intervention found in the 1990s to positively impact adolescent sexual behavior. That smaller randomized controlled trial showed evidence that compared to a control group, those who were randomized to view *What Could You Do?* were significantly more likely to be abstinent and have fewer condom failures at three months post-intervention and less likely to have Chlamydia at six months post-intervention (Downs, Murray, Bruine de Bruin, Penrose, Palmgren & Fischhoff, 2004).

However, this evaluation team, using a randomized controlled trial design with an enrolled sample of over 1,300 adolescent females in three Mid-Atlantic and Midwestern states, found no general effects of *Seventeen Days* on reducing risky sexual behavior at the six-month post-randomization assessment. Adolescents who were randomized to view *Seventeen Days* were no more likely than those who were randomized to view a driving skills video to report abstinence or using a condom consistently and correctly at every sexual encounter in the three months prior to their six-month post-randomization assessment. Unlike the *What Could You Do?* evaluation, this study did not find a statistically significant effect on safer sexual behavior at the three-month post-randomization assessment. Our study did not find significant evidence of impacts on abstinence at the three-month post-randomization assessment. There were also no significant impacts on either pregnancy or STI rates at the six-month post-randomization assessment.

The impact findings may be influenced by several significant limitations of the study. One limitation was the poor quality behavioral data received from the survey's Timeline Followback

Calendar tool that provided the data for the primary research question. The intention of using this type of instrument was to facilitate detailed recall and provide rich, nuanced behavioral data. However, the tool was time-consuming for participants to complete and that burden led to a high rate of non-completion on some key behavioral survey questions. There is no existing literature on using an electronic version of the Timeline Followback Calendar and unfortunately, due to substantial delays in completing the programming before the evaluation needed to start, there was not adequate time to pilot test the electronic calendar in sufficient depth to discover these shortcomings. Additionally, a few programming errors led to missing data. For example, during a period of several months early in the study, baseline responses were overwritten when participants returned to the survey to complete their three-month post-randomization survey. Another programming error occurred during a system update, in which all of the behavioral data reported on protection was deleted from the system. This error affected both treatment groups equally. Both of these errors resulted in an unexpected amount of missing baseline data in the analyses.

Secondly, implementation challenges impacted the ability of participants to view the intervention as originally intended, within wait time at clinics. Every clinic in the study had some level of technological challenge such as poor and unreliable Internet connections and in many clinics the typical wait time for care was shorter than expected. Many participants were not able to view the video during their clinic appointment because all of their wait time was consumed with screening, enrollment, consent, and baseline data collection.

It is logical to hypothesize that an increase in safe sexual behavior may be dose-dependent, that is, the more video material a participant views, the more likely she will show effects of the intervention. The relationship between dose, viewing patterns, and participant outcomes will be

explored in future analyses. Other planned analyses will explore participant characteristics such as age, sexual experience at baseline, relationship status, and others; viewing choices and dose; and the program's impacts on sexual health knowledge and self-efficacy for negotiation safer sexual behavior.

VI. References

Hamilton B, Martin J, Osterman M, Curtin S. Births: Preliminary Data for 2014. *National Vital Statistics Reports*. 2015; 64, 1-18.

Sedgh G, Finer L, Bankole A, Eilers M, Singh S. Adolescent pregnancy, birth, and abortion rates across countries: levels and recent trends. *J Adolesc Health*. 2015; 56(2): 223-30.

Kost K. & Henshaw S. *U.S. Teenage Pregnancies, Births and Abortions 2008: National Trends by Race and Ethnicity*. *Guttmacher Institute*. 2012.

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

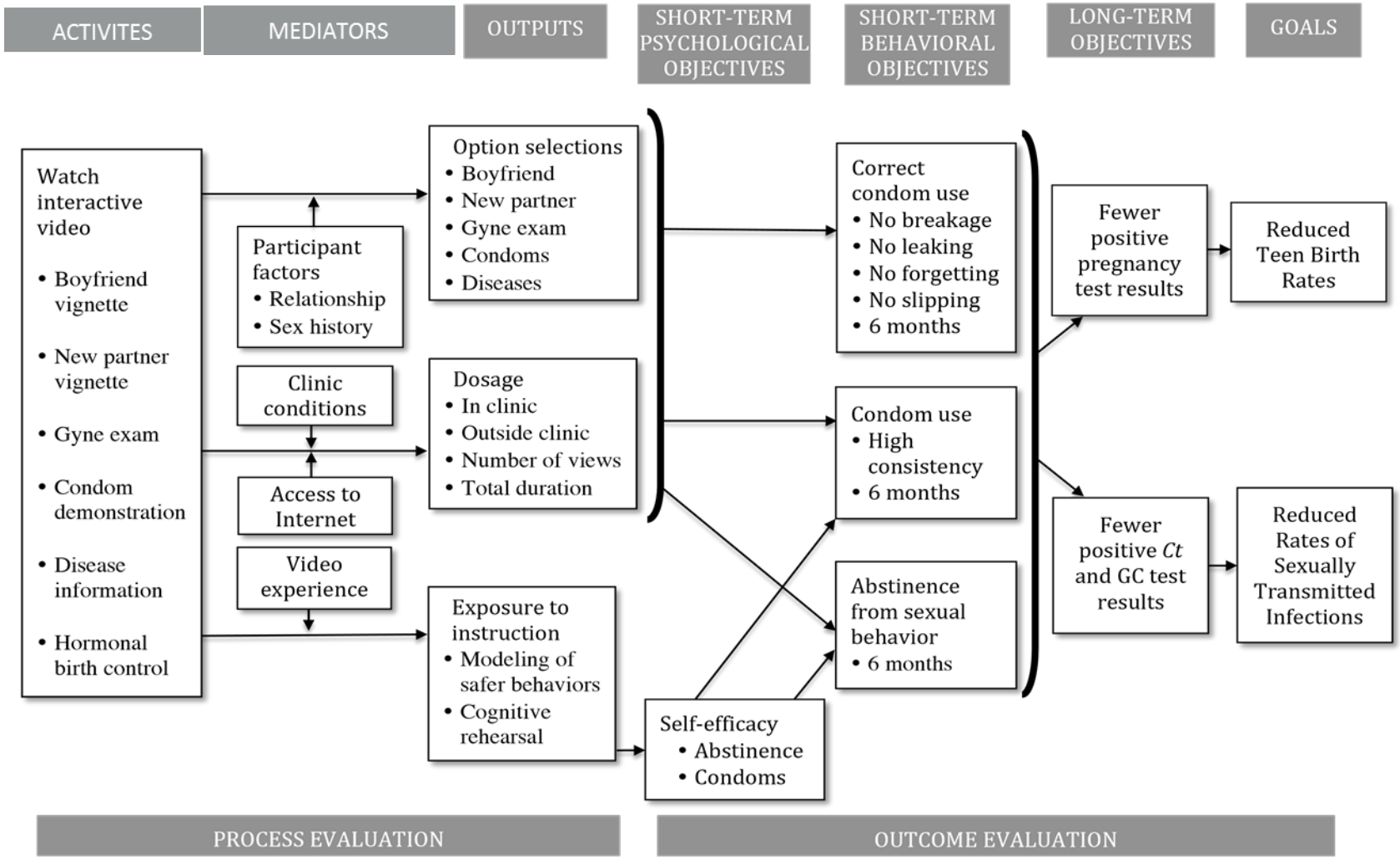
Downs JS, Murray PJ, Bruine de Bruin W, Penrose J, Palmgren C, Fischhoff B. Interactive video behavioral intervention to reduce adolescent females' STD risk: A randomized controlled trial. *Social Science & Medicine*. 2004; 59: 1561-1572.

Mathematica Policy Research & Child Trends. Identifying programs that impact teen pregnancy, sexually transmitted infections, and associated sexual risk behaviors. 2010. Review Protocol Version 1.0.

Weinhardt LS, Carey MP, Maisto SA, Carey KB, Cohen MM, Wickramasinghe SM. Reliability of the Timeline Followback Sexual Behavior Interview. *Annals of Behavioral Medicine.* 1998; 20: 25-30.

Sobell LC, Sobell MB. Timeline Follow-Back. *Measuring Alcohol Consumption.* 1992: 41-72.

Appendix A: Intervention Logic Model



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: How often was the video offered?	Intervention viewing data shows number of times each participant logged in and the viewing choices made by participants	Data are collected electronically each time participants logs in; intervention was available to participants for 6 months of unlimited access	Data are collected electronically in real time as participant views video intervention and completes electronic data collection tools
Adherence: What and how much was received?	Intervention viewing data shows number of views, proportion of total material viewed, and viewing choices made by participants	Data are collected electronically each time participants logs in; amount and content of material viewed varied by participant, after core dose was completed	Data are collected electronically in real time as participant views video intervention and completes electronic data collection tools
Adherence: What content was delivered to youth?	Intervention viewing data show if core material was viewed and the additional content viewed by participants beyond core material	Data are collected electronically each time participants logs in; amount and content of material viewed varied by participant, after core dose was completed	Data are collected electronically in real time as participant views video intervention and completes electronic data collection tools
Counterfactual: Experiences of comparison condition	Intervention viewing data shows number of views and viewing choices made by participants Intervention viewing data show if core material was viewed and the additional content viewed by participants beyond core material	Data are collected electronically each time participants logs in; amount and content of material viewed varied by participant, after core dose was completed	Data are collected electronically in real time as participant views video intervention and completes electronic data collection tools

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	Time to complete viewing of intervention in clinic	1-3 observations were conducted at a sample of clinics over the recruitment period	Program staff and Evaluation staff Oral and written reports by program staff
	Technology issues in some clinics, e.g., poor Internet access	Monthly team meetings with recruiters, program support calls with the Office of Adolescent Health, and regular information sharing with project team; Recruiter interviews	Data are collected electronically in real time as participant views video intervention and completes electronic data collection tools
	Data on the number of participants who viewed alternate video due to programming glitch	Data are collected electronically each time participant logs in	

Appendix C: Study sample

Table C.1. Youth sample sizes by intervention status

Number of youth	Time Period	Total sample size	Intervention sample size	Comparison sample size	Total response rate	Intervention response rate	Comparison response rate
Assigned to condition	Immediately pre-intervention	1,317	653	664			
Contributed a baseline survey	Prior to randomization	1,263	629	634	.96	.96	.96
Contributed a follow-up survey	3 months post-randomization	685	342	343	.52	.52	.52
Contributed a follow-up survey	6 months post-randomization	674	334	340	.51	.51	.51
Contributed a test kit	6 months post-randomization	563	280	283	.43	.45	.45

Note: Due to a programming glitch, some baseline data was over-written, and therefore unavailable, for the outcome variables of interest.

Appendix D: Impact Model Specification

Impact models for outcomes at 6-month follow-up

$$\begin{aligned} Y = {} & \beta_0 + \beta_T \text{Condition} + \beta_{D1} \text{Baseline (missing)} + \beta_{x1} \text{Baseline} \\ & + \beta_{D2} \text{Ethnicity (missing)} + \beta_{x2} \text{Ethnicity} \\ & + \beta_{D3} \text{Race (missing)} + \beta_{x3} \text{Race} + \beta_{x4} \text{Clinic} + \beta_{D4} \text{Age (missing)} + \beta_{D5} \text{Age} \end{aligned}$$

This is the basic model used in all analysis. Y represents the dependent variable of interest, β_0 is the intercept of the model; β_T is the regression coefficient for the treatment effect; β_{D1} is the regression coefficient for the dummy variable that indicates whether the value of the outcome variable of interest is missing; β_{x1} is the coefficient for the baseline outcome variable of interest; β_{D2} is the coefficient for dummy variable that indicates whether participants' ethnicity is missing; β_{x2} is the coefficient for the participants who are Hispanic; β_{D3} is the coefficient for dummy variable that indicates whether the participants' race is missing; β_{x3} are the coefficients for the host of dummy variables that represent each racial group; β_{x4} are the coefficients for the dummy variables that represent each clinic; β_{D4} is the coefficient for the dummy variable that indicates whether participants' age is missing; β_{x5} is the coefficient for the participants' age.

Model 1:

$$\text{Abstinence (6 months)} = \beta_0 + \beta_T \text{Condition}$$

The basic model only included condition as the predictor. In this case, Condition is a dichotomous variable indicating whether the participants were assigned to the intervention or comparison group.

Model 2:

Abstinence (6 months)

$$\begin{aligned} = {} & \beta_0 + \beta_T \text{Condition} + \beta_{D1} \text{Abstinence baseline (missing)} + \beta_{x1} \text{Baseline} \\ & + \beta_{X1} \text{Abstinence baseline} + \beta_{D2} \text{Ethnicity (missing)} + \beta_{x2} \text{Ethnicity} \\ & + \beta_{D3} \text{Race (missing)} + \beta_{x3\text{-}x6} \text{Race} + \beta_{x7\text{-}x12} \text{Clinic} + \beta_{D4} \text{Age (missing)} + \beta_{B13} \text{Age} \end{aligned}$$

In the full model, baseline covariates are added to the model. Abstinence baseline(missing) is a dichotomous variable indicating whether the participants answered the abstinence question at baseline; Abstinence baseline is a dichotomous variable indicating whether the participants self-reported to be abstinent at baseline; Ethnicity(missing) is a dichotomous variable indicating whether the participants reported their ethnicity; Ethnicity is a dichotomous variable that

represents whether the participants are Hispanic; Race(missing) is a dichotomous variable indicating whether racial information is missing for the participants; Race represents four dummy variables with each dummy variable represents one racial group and "more than one race" is the reference group; Clinic represents sixteen dummy variables, where each dummy variable represents one clinic site. The clinic, "WV-Randolph", is the reference group in this set of variables; Age(missing) is a dichotomous variable indicating whether participants reported their age; and finally, Age is a continuous variable of the participants' age reported in years.

Model 1:

$$\text{Safe sexual behavior (6 months)} = \beta_0 + \beta_T \text{Condition}$$

First model in each set of analysis has this same form. Only exception is safe sexual behavior, where a missing dummy is included for one set of analysis to account for the "No, assumed" category.

Model 2:

Safe sexual behavior (6 months)

$$\begin{aligned} &= \beta_0 + \beta_T \text{Condition} + \beta_{D1} \text{Safe sexual behavior baseline (missing)} \\ &\quad + \beta_{X1} \text{Safe sexual behavior baseline} + \beta_{D2} \text{Ethnicity (missing)} \\ &\quad + \beta_{x2} \text{Ethnicity} + \beta_{D3} \text{Race (missing)} + \beta_{x3\text{-}x6} \text{Race} \\ &\quad + \beta_{x7\text{-}x12} \text{Clinic} + \beta_{D4} \text{Age (missing)} + \beta_{B13} \text{Age} \end{aligned}$$

In this model, Safe sexual behavior baseline(missing) is a dichotomous variable indicating whether the participants reported their sexual behavior at baseline; Safe sexual behavior baseline is a dichotomous variable indicating whether the participants reported to engage in safe sexual behavior; Ethnicity(missing) is a dichotomous variable indicating whether the participants reported their ethnicity; Ethnicity is a dichotomous variable that represents whether the participants are Hispanic; Race(missing) is a dichotomous variable indicating whether racial information is missing for the participants; Race represents four dummy variables with each dummy variable represents one racial group and "more than one race" is the reference group; Clinic represents sixteen dummy variables, where each dummy variable represents one clinic site. The clinic, "WV-Randolph", is the reference group in this set of variables; Age(missing) is a dichotomous variable indicating whether participants reported their age; and finally, Age is a continuous variable of the participants' age reported in years.

Model 1:

$$\text{Pregnancy (6 months)} = \beta_0 + \beta_T Condition$$

Only a dichotomous variable that indicates whether the participants were assigned to the intervention or comparison group in included in the base model.

Model 2:

Pregnancy (6 months)

$$\begin{aligned} &= \beta_0 + \beta_T \text{Condition} + \beta_{D1}\ \text{Pregancy history} + \beta_{D2}\ \text{Ethnicity (missing)} \\ &\quad + \beta_{X2}\text{Ethnicity} + \beta_{D3}\text{Race (missing)} + \beta_{x3\text{-}x6}\text{Race} \\ &\quad + \beta_{x7\text{-}x12}\text{Clinic} + \beta_{D4}\text{Age (missing)} + \beta_{B13}\text{Age} \end{aligned}$$

In this model, Pregnancy history is a dichotomous variable indicating whether the participants were ever pregnant; Ethnicity(missing) is a dichotomous variable indicating whether the participants reported their ethnicity; Ethnicity is a dichotomous variable that represents whether the participants are Hispanic; Race(missing) is a dichotomous variable indicating whether racial information is missing for the participants; Race represents four dummy variables with each dummy variable represents one racial group and "more than one race" is the reference group; Clinic represents sixteen dummy variables, where each dummy variable represents one clinic site. The clinic, "WV-Randolph", is the reference group in this set of variables; Age(missing) is a dichotomous variable indicating whether participants reported their age; and finally, Age is a continuous variable of the participants' age reported in years.

Model 1:

$$\text{Infection (6 months)} = \beta_0 + \beta_T \text{Condition}$$

The base model only includes a dichotomous variable that indicates whether the participants were assigned to the intervention or comparison group.

Model 2:

Infection (6 months)

$$\begin{aligned} &= \beta_0 + \beta_T \text{Condition} + \beta_{D1}\text{Infection baseline (missing)} \\ &\quad + \beta_{X1}\text{Infection baseline} + \beta_{D2}\text{Ethnicity (missing)} + \beta_{x2}\text{Ethnicity} \\ &\quad + \beta_{D3}\text{Race (missing)} + \beta_{x3\text{-}x6}\text{Race} + \beta_{x7\text{-}x12}\text{Clinic} + \beta_{D4}\text{Age (missing)} \\ &\quad + \beta_{B13}\text{Age} \end{aligned}$$

In this model, Infection baseline(missing) is a dichotomous variable indicating whether the participants' infection history at baseline is missing; Infection baseline is a dichotomous variable indicating whether the participants were found to be infected; Ethnicity(missing) is a

dichotomous variable indicating whether the participants reported their ethnicity; Ethnicity is a dichotomous variable that represents whether the participants are Hispanic; Race(missing) is a dichotomous variable indicating whether racial information is missing for the participants; Race represents four dummy variables with each dummy variable represents one racial group and "more than one race" is the reference group; Clinic represents sixteen dummy variables, where each dummy variable represents one clinic site. The clinic, "WV-Randolph", is the reference group in this set of variables; Age(missing) is a dichotomous variable indicating whether participants reported their age; and finally, Age is a continuous variable of the participants' age reported in years.

Appendix E: Sensitivity analyses

To test whether the results presented in the benchmark analysis were sensitive to researcher decisions about how data were cleaned and analyzed, we conducted three sensitivity analyses. The first sensitivity analysis tests the model without controlling for baseline covariates. The subsequent sensitivity analyses were conducted on the safe sexual behavior outcomes constructed from the Timeline Followback Calendar. These sensitivity analyses test the decisions made about handling missing calendar data; if participants did not complete a six month test kit, or provide a response to the question, "Have you had sex in the last three months?" on the calendar, they were not included in the analytic samples for those benchmark analyses, therefore, these sensitivity analyses were not conducted on those outcomes. The second sensitivity analysis includes only participants who provided a protection response on the Timeline Followback Calendar. In the third sensitivity analysis, a dummy variable was created for imputed protection responses on the Timeline Followback Calendar and added to the regression model. The results of the sensitivity analyses are presented in Tables E.1 and E.2.

Table E.1. Sensitivity of impact analyses using data from Timeline Followback Calendar to address the primary research question

Intervention compared with comparison	Benchmark approach difference	Benchmark approach *p*-value	No baseline covariates difference	No baseline covariates *p*-value	No imputed protection responses difference	No imputed protection responses *p*-value	Dummy variable for imputed responses difference	Dummy variable for imputed responses *p*-value
Safe Sexual Behavior	0.007	0.873	0.011	0.767	0.007	0.876	0.001	0.975

Source: Timeline Followback Calendar administered 6 months post-randomization.
Notes: Safe sexual behavior is defined as either abstinent from sexual activity or no reports of vaginal sexual events unprotected by a condom in the last 3 months.

Table E.2. Sensitivity of impact analyses using data from Timeline Followback Calendar and Clinic Test Kit to address the secondary research questions

Intervention compared with comparison	Benchmark approach difference	Benchmark approach *p*-value	No baseline covariates difference	No baseline covariates *p*-value	No imputed protection responses difference	No imputed protection responses *p*-value	Dummy variable for imputed responses difference	Dummy variable for imputed responses *p*-value
Safe Sexual Behavior at 3 months	0.029	0.402	0.019	0.615	0.053	0.176	0.047	0.130
Abstinence at 3 months	0.001	0.982	0.005	0.864				
Abstinence at 6 months	0.008	0.790	0.014	0.652				
Pregnancy Test Result	0.004	0.723	0.003	0.750				
Infection Test Result	0.005	0.817	0.012	0.615				

Source: Timeline Followback Calendar administered at 3 or 6 months post-randomization. Clinic Test Kit administered at 6 months post-randomization.

Notes: Safe sexual behavior is defined as either abstinent from sexual activity or no reports of vaginal sexual events unprotected by a condom in the last 3 months

Appendix F: Implementation evaluation methods

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

Implementation element	Methods used to address each implementation element
Adherence: How often was the video offered?	Online curriculum; full dosage of 2.5 hours of programming offered to all participants; intervention was available to participants for 6 months of unlimited access
Adherence: What and how much was received?	Percentage of baseline material each participant viewed (calculated as the amount of material viewed divided by all possible baseline material) Percentage of participants who completed baseline viewing (calculated as the number who completed baseline divided by the total number of participants) Percentage of additional material each participant viewed Percentage of participants who chose to view additional video beyond baseline dosage (calculated as number of those who viewed additional video divided by the total number of participants)
Adherence: What content was delivered to youth?	Description of the content contained in the baseline dosage and description of the type of content most frequently viewed during additional log-ins
Counterfactual: Experiences of counterfactual condition	Percentage of baseline material each participant viewed (calculated as the amount of material viewed divided by all possible baseline material) Percentage of participants who completed baseline viewing (calculated as the number who completed baseline divided by the total number of participants Percentage of additional material each participant viewed Percentage of participants who chose to view additional video beyond baseline dosage (calculated as number of those who viewed additional video divided by the total number of participants)
Context: External events affecting implementation	Number of clinics where insufficient waiting time (i.e., time to complete the intervention as intended) Number of clinics where technology issues prohibited or hindered participation Number of participants who viewed alternate video due to programming glitch