Findings from an Innovative Teen Pregnancy Prevention Program

# Evaluation of Will Power/Won't Power in Los Angeles County

Final Impact Report for

Volunteers of America – Girls Inc. of Greater Los Angeles

October 31, 2015

Submitted with revisions on May 6, 2016

Prepared by

Advanced Empirical Solutions, LLC

Advanced Empirical Solutions on the Impact of a Teen Pregnancy Prevention Program
in Los Angeles County School Districts, Evaluation of Will Power/Won't Power in Los
Angeles County (Advanced Empirical Solutions, 2015)

#### Acknowledgements:

We are grateful for the steadfast and meaningful contributions of our VOALA Girls Inc. partners, including CFO Patrice Louie, Internal Director of Research and Evaluation Suzanne Markoe Hayes, Program Manager Cristina Ramirez and Program Innovations Manager Cindi Dietrich. The depth knowledge of our target youth population and the commitment to implementing with culturally responsive attention to Los Angeles girls has been invaluable to the investigation. We appreciate your praxis and community. We would also like to extend deep gratitude to the institutions that repeatedly and graciously allowed us to implement the afterschool program in their districts and sites. This includes teams in Downey, Lynwood, Culver City and Los Angeles Unified School Districts. We thank Heartland IRB for their guidance in implementing an ethical study. Last, we are indebted to the keen leadership of Program Manager Sabrina Chapple and Mathematica Consultants Jacqueline Berman and Russell Cole who steadied this journey.

This publication was prepared under Grant Number TP2AH000019-01-00 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.

#### Note:

In the following report, Volunteers of America Los Angeles (VOALA) and evaluation partners, Advanced Empirical Solutions (AES), examined the impact of a teenage pregnancy prevention initiative on a sample of middle schools students in Los Angeles County.

The initial design included both high school and middle school samples; however, unmanageable attrition among high school participants posed challenges to collecting meaningful data. With OAH and Mathematica approval, VOALA withdrew its high school sample from the study. The opportunity to examine rates of sexual activity onset and pregnancy diminished with this withdrawal; the incidence of sexual activity and pregnancy rates among middle school students in Los Angeles was negligible prior to and during the OAH study period. The absence of rates required that VOALA redirect attention from sexual activity onset and pregnancy incidence to feasible measures for the middle school sample – knowledge outcomes.

Per the OAH contract, VOALA reports outcomes on the original OAH primary questions, which have scant data from the middle school sample, and knowledge outcomes.

# Contents

Introduction.	6
Introduction and Study Overview	6
Primary Research Question(s)	7
Secondary Research Question(s)	7
Intervention and Counterfactual Programming	8
Description of Will Power/Won't Power Program as Intended	8
Description of Equal Earners, Savvy Spenders Counterfactual Condition	9
Study Design	10
Sample Recruitment	10
Study Design	12
Data Collection	13
Impact Evaluation	13
Implementation Evaluation	14
Outcomes for Impact Analyses	15
Study Sample	18
Pre-survey Equivalence	19
Methods	21
Impact Evaluation	21
Implementation Evaluation	22
Study Findings	23
Impact Study Findings	23
Implementation Study Findings	26
Conclusions.	30
References.	32
Annendives	33

#### INTRODUCTION

## **Introduction and Study Overview**

Teen pregnancy rates in California are gradually declining with the exception of a slight upturn in Latino teen rates (Appendix A). Los Angeles County is particularly plagued by teen pregnancy risk factors, including high sexual activity and unprotected sex rates, lack of access to pregnancy prevention resources, and sexually transmitted diseases among teens (Advocates for Youth, 2009). County data collected through the Youth Risk Behavior Surveillance Survey (YRBSS) revealed that 46% of high school students had engaged in sexual intercourse—7% before the age of 13. Twelve percent of these youth had intercourse with four or more persons prior to the completing the survey, and 32% reported being sexually active at time of the survey. Further, 34% of youth responding to the YRBSS did not use a condom during their last sexual intercourse, and 20% drank alcohol or used drugs before that encounter (Eaton, Kann, Kinchen, Shanklin, & Ross et al., 2007).

Sexually transmitted infections were high among Los Angeles County youth. Recent chlamydia rates for youth ages 10-14 was 817.4 per 100,000 infections in 2008, up 8% since 2006. Statewide primary and secondary syphilis cases for youth ages 15-24 are 6.4 per 100,000 compared to the county's 10 per 100,000 (Centers for Disease Control and Prevention [CDC], 2009). Birth rates among teens in communities deemed high-need<sup>1</sup> by the California Public Health Department (2009), including Hollywood, Boyle Heights, West LA, East LA and South LA, were 1.5- to 2.5-times higher than state and national birth rates.

Girls Inc. of Greater Los Angeles (a Volunteers of America Los Angeles [VOALA] program) implemented an innovative afterschool program designed to prevent teenage pregnancy aligned with the Office of Adolescent Health/Family Youth Services Bureau Tier 2 grant goals. Tier 2 grants were

<sup>&</sup>lt;sup>1</sup> Need refers to the prevalence of health, safety and education services in a geographic region identified by public health agencies. High need refers to a higher prevalence of unmet needs.

awarded to demonstrate and improve on existing programs or to formally explore new programs designed for "youth ages 10-19, with a particular interest in reaching high-risk, vulnerable, and culturally under-represented youth populations" (Department of Health and Human Services, 2016). This evaluation examined the implementation and impact of an abstinence-plus pregnancy prevention curriculum (Will Power/Won't Power; WPWP) on girls' sexual health outcomes compared to any impact of an economic literacy counterfactual curriculum (Equal Earners, Savvy Spenders; EESS) using a randomized controlled trial. The program targets girls attending Title 1<sup>2</sup> middle schools in LA County with large Latino student populations.

#### **Primary Research Question(s)**

The current evaluation examined the long-term impacts of WPWP on sexual activity onset and pregnancy incidence. The two primary research questions were:

- 1. What is the impact of the WPWP curriculum relative to the EESS curriculum on participants' sexual activity onset one year after the end of the program?
- 2. What is the impact of the WPWP curriculum relative to the EESS curriculum on participants' pregnancy incidence one year after the end of the program?

#### **Secondary Research Question(s)**

The secondary research questions (1) evaluate the primary research questions at the end of the program and six months post-program, and (2) explore participants' knowledge and perceptions at all three post-survey points.

1. What is the impact of the WPWP curriculum relative to EESS curriculum on participants' sexual health knowledge and perceived barriers to sexual health immediately following the program, six months after the end of the program, and one year after the end of the program?

<sup>&</sup>lt;sup>2</sup> Title 1 is a provision of the Elementary and Secondary Education Act of 1965 that authorizes additional funding for schools with high prevalence of students from low-income homes, children at-risk of maltreatment and neglect and immigrant youth. To receive the funding, schools must show that at least 40% of their student body meets the low-income criterion as defined by the United States Census. See Carmichael (1997) for elaboration.

2. What is the impact of the WPWP curriculum relative to EESS curriculum on participants' sexual intentions immediately following the program, six months after the end of the program, and one year after the end of the program?

#### INTERVENTION AND COUNTERFACTUAL PROGRAMMING

# Description of Will Power/Won't Power Program as Intended

Will Power/Won't Power (WPWP) is an interactive 10-course curriculum wherein "girls build skills and strategies for dealing with sexual situations as they enter the most pressure-sensitive adolescent years, while also receiving medically accurate information" (Girls Inc., 2016). The courses include female reproductive anatomy, assertiveness, peer support and/or independence in good decision-making, parents as resources on sexual health information, hygiene and personal values among other lessons immediately relevant to girls' making sound choices about their bodies and their sexual behaviors.

Girls in a WPWP classroom receive 11 weekly afterschool sessions. Each session is intended to last approximately 1.5 hours. Of the 11 sessions, 10 are curricular content and the last session is a parent/guardian and daughter workshop wherein girls share what they have learned with their accompanying adult(s).

Program Specialists hired and trained by the Girls Inc. program manager implemented WPWP in classrooms. Specialists were women ranging in age from their early 20s to mid 30s experienced in successful engagement and communication with teens. Twice during the school year/implementation year, each specialist completed a total of 12 WPWP-specific training hours over three days (four hours daily). Training included in-depth instruction on curriculum content and tools, presentation skills and behavior management strategies. To note, specialists were trained in the counterfactual curriculum also, receiving 10 hours of instruction over three days twice annually (3.5 hours for two days and three hours

on one day). The Program Manager randomized Program Specialists to classrooms, requiring that specialists be well versed in both curricula.

## Description of Equal Earners, Savvy Spenders Counterfactual Condition

VOALA selected a counterfactual that did not provide any health information to ensure that girls in that classroom did not experience curricular contamination, meaning that girls in the counterfactual would have no interaction with health, sexual health or reproductive health. While the organization could not control what information girls learned about their bodies in regular school courses or afterschool programs hosted by other nonprofits, VOALA could at least isolate all health instruction to the intervention classrooms among its participants. Girls in counterfactual classrooms would still receive meaningful skills that facilitate individual decision-making, through the *Equal Earners, Savvy Spenders* (EESS) curriculum, however the focus would be finances—a topic that is not instructed often in the county's schools but is immediately practical and relevant to girls' autonomies.

The EESS curriculum is designed to teach girls about money and the economy, including how to manage, invest, and save money and how to help others through philanthropy. The curriculum imparts skills important to being smart about finances and to becoming economically independent adults. Core components give girls a foundation for an economically independent adulthood, an understanding of key economic concepts at the individual, family, community, national and global levels, and girls' impact on the system.

Program Specialists implement EESS across 10 weekly afterschool sessions. Each session lasts approximately 1.5 hours. Specialists are women ranging in age from their early 20s to mid 30s experienced in successfully engaging and communicating with teens. Their training included in-depth instruction on content, presentation, tools, and behavior management. Each specialist completed the 3-day EESS training twice during the school year/implementation year. Specialists received training in

both EESS curriculum and WPWP curriculum to maximize their preparedness when randomized to a classroom

# **Study Design**

# **Sample Recruitment**

Girls in 6th and 7th grades across 12 non-specialty (pregnant and parenting programs, probation programs, charter) public schools with a command of the English language were eligible for the study.

Recruitment and program implementation began in spring 2012 and ended in spring 2015. In coordination with school leadership, VOALA outreached to girls using the following approaches:

- 1. The program team shared a flyer for school staff to include in parent packets as part of the beginning of school year process.
- 2. School staff made verbal announcements, and teen coordinators distributed fliers during nutrition, lunch, and physical education.
- 3. Teachers, counselors and administrators referred girls to the program.
- 4. Girls Inc. staff hosted informational meetings for parents/caregivers and daughters.
- 5. Staff introduced attendees to the program, potential participation benefits and participation incentives (field trips, gift certificates, small gifts).

Girls could sign-up individually or as a pair during the recruitment period. VOALA developed a *Girl+1* strategy to allow friends simultaneous enrollment and, eventually, randomization as one unit. The strategy was designed to retain girls by allowing two friends to move through a semester of programming in the same classroom.

Participation in the programming was voluntary. To be eligible for participation and to enter either type of classrooms, girls must meet three preliminary criteria:

- 1. Girls must return a signed parent/guardian informed consent form. Girls who were 18 years of age (legal adults in California) must also submit a signed youth assent form.
- 2. Girls must speak and understand the English language as Program Specialists implemented curricula in English.
- 3. Girls must not have previously participated in WPWP or EESS.

To limit low attendance, the Program Manager randomized participants only after they submitted a verified consent form and attended an introductory session beginning with Cohort 3.

#### Consent

The parent/guardian consent form to be completed for each girl receiving programming from Girls Inc. during the grant period articulated that Girls Inc. would offer two afterschool sessions — "one on reproductive health and a different one on economic literary." The consent form clarified that VOALA would randomly assign girls to one condition and that girls and parents may not choose.

Parents/guardians received an alert that their girls "might receive information about reproductive health." Parents could call the Girls Inc. office with additional questions using contact information provided on the form.

The consent form was comprehensive in its description of randomization, girls' voluntary participation in either curricula, data collection activities in which girls could choose to take part, and a mandated reporting notice. All girls who entered the classrooms were required to submit parent/guardian consent forms because of the study's nature—random assignment might place girls in a classroom where reproductive and sexual health were taught. Girls who opted into the study were invited to participate in the survey and focus group components and were included in a study retention incentive structure. Girls who did not opt into the study only received curriculum; while these girls remained in the classrooms, we did not collect survey or focus group data. We note this in the study as missing data.

#### Randomization

During Cohorts 1 and 2 (a and b), the Program Manager randomized girls immediately upon receipt of their signed parent/guardian consent forms and assent forms. To reduce the likelihood of low attendance, the Program Manger randomized girls in Cohorts 3 and 4 after they submitted the required forms and attended a pre-session. VOALA designed the pre-sessions to be social spaces where girls

played games, ate snacks and heard an explanation of the upcoming sessions. The Program Manager used this pre-session as an orientation and an initial retention scan to more clearly identify which girls would be likely and available to attend the afterschool sessions. Girls who attended the pre-session were unlikely to be involved in competing activities (sports, tutoring clubs) that took place concurrently with the intervention and counterfactual sessions; therefore, these girls could participate fully in the intervention and counterfactual sessions.

Late enrollment. VOALA gave latitude to girls who missed the pre-session for reasons other than competing activities (e.g., they submitted late consent forms or were absent from school). VOALA allowed girls to join up to three weeks following the start of Girls Inc. programming—the equivalent of three sessions (one introductory session and two content sessions). Upon a late enrollee's arrival, the Program Manager immediately randomized the student. If the student missed the pre-survey period, then she was still welcome to complete subsequent surveys and participate in focus groups.

# **Study Design**

The evaluation is a randomized controlled trial. For the 5-year duration of the grant, Girls Inc. discontinued their typical programming and adopted the WPWP or EESS. The student is the unit of randomization. The Program Manager randomly assigned girls to classrooms by adding names to each line of a site-specific randomization spreadsheet developed by the evaluation team. Individual lines were pre-assigned an ID number and condition (intervention or counterfactual) for individual participants. The manager entered one girl's name on each consecutive line, filling a page for one school without leaving empty lines. Ten lines on each spreadsheet were coupled into five pairs for girls who enrolled with a friend using a *Girl+1* clustering strategy that enabled girls to be assigned to condition together to increase enrollment and retention. Girls in these pairs have different ID numbers but share the same condition. In analyses, one randomly selected girl from each cluster was included.

Randomization continues on a rolling basis as consent documents are received and approved until the third session. In total, 500 girls received an intervention classroom (WPWP) assignment and 498 received a counterfactual (EESS) classroom assignment (Table 1).

Table 1. Sample Sizes at the Time of Random Assignment

	Intervention	Counterfactual	Total Randomized
Cohort 1	52	43	95
Cohort 2 a+b	129	126	255
Cohort 3	182	196	378
Cohort 4	137	133	270
Total	500	498	998

#### **Data Collection**

The evaluation team collected data using three means throughout the school year: surveys, focus group responses and observations. Program specialists, responsible for teaching participants, collected attendance data during each session that the evaluation team later used to calculate any dosage impact.

#### Impact Evaluation

In each classroom, data collectors proctored paper surveys to participants in intervention and counterfactual groups at four times: prior to the first session (pre-survey), immediately following the end of program, six months following the end of the program and 12 months following the end of the program. Pre-survey took place in one classroom with all girls prior to their division into separate classrooms. After the program ended, girls completed their first post-survey in their designated classrooms (intervention or counterfactual). The evaluation team collected the remaining two post-surveys from girls via ground mail at six and 12 months post-program completion. The evaluation team initially did not incentivize survey completion. However, VOALA began to reward girls who returned their 6- and 12-month post-surveys by mail with \$20 Forever 21 and Target gift cards for Cohorts 2a and 2b, which improved response rates.

## Implementation Evaluation

The implementation evaluation focused on four objectives that are deemed critical to program implementation quality. These objectives include: 1) program adherence, 2) program quality, 3) counterfactual programing, and 4) program context. For program adherence and counterfactual programing, data regarding session attendance (i.e., how many sessions were offered, how many sessions were received) and session content (i.e., how many topics were delivered, who delivered the material) were collected. Quality, programming and context were addressed through focus groups and observations.

The evaluation team designed focus group questions to supplement survey items, clarify how cultural and religious practices might shape girls' knowledge about sexual activity and pregnancy, and provide additional insight into the quality of staff-participant interactions and the quality of girls' engagement with the program. Data collectors conducted two types of focus groups with purposefully selected participants at selected schools on the last day of the program. The diverse groups (DG) included a range of girls in the cohorts who were interested in sharing their perspectives. Data collectors selected DG participants based on their ages, races and ethnicities and grades to capture diverse responses. The target groups (TG) only included girls of one particular race or ethnic background (e.g., Asian American students) to more deeply explore the impact of culture and religions on health knowledge, if any. A total of 69 girls participated in the focus groups, which took place the week following each cohort's end (see Table 2).

**Table 2. Focus Group Participation Summary** 

	Cohort	Date	Participant Count	Group Type
1.	1	December 7, 2011	6	DG
2.	1	December 9, 2011	6	DG
3.	2	May 9, 2012	5	DG
4.	2	March 15, 2013	2	TG
5.	3	January 14, 2014	3	DG
6.	3	January 14, 2014	6	TG
7.	3	January 21, 2014	4	TG
8.	3	January 29, 2014	8	TG
9.	4	May 16, 2014	7	DG
10.	4	May 27, 2014	2	DG
11.	4	Date unknown	5	DG
			69	7 DG, 4 TG

Additionally, observations took place throughout each semester. Data collectors observed classrooms to capture implementation quality, curriculum fidelity and participant engagement. Data collectors initially observed randomly selected intervention and counterfactual classrooms during Cohorts 1, 2a and 2b. Beginning with Cohort 3, data collectors observed one specific WPWP session at all sites, resulting in consistent data collection and engagement with 10% of all intervention classes as clarified by the Office of Adolescent Health's guidelines. The implementation evaluation measures, data collection plan, and timeline are more fully described in Appendix B.

# **Outcomes for Impact Analyses**

The outcome survey is a composite of items from four sources. Survey items included those required by the Office of Adolescent Health, select items from the California Health Interview Survey select items from Adolescent Health studies, and items about girls' intent related to sexual activity. VOALA and AES selected items to respond to both primary and secondary research questions. Tables 3.1 and 3.2 define the outcomes being examined in the primary and secondary research questions and how they were constructed.

Table 3.1. Behavioral Outcomes Used for Primary Impact Analyses

Outcome Name	Outcome Description	Timing of Measure Relative to Program			
Sexual Activity Onset	The variable is a yes/no measure of whether a person has ever had sexual intercourse. The measure is taken directly from the following item on the survey:	Pre-Survey and one year after the completion of the program			
	"Have you ever had sexual intercourse?"				
	The variable is constructed as a dummy variable where respondents who respond yes they have had sex are coded as 1 and all others are coded as 0.				
Pregnancy Incidence	The variable is a yes/no/l don't know measure of whether a person has ever been pregnant or gotten someone pregnant. The measure is taken directly from the following item on the survey:	Pre-Survey and one year after the completion of the program			
	<ul> <li>"To the best of your knowledge, have you ever been pregnant or gotten someone pregnant, even if no child was born?"</li> </ul>				
	The variable is constructed as a dummy variable with where respondents who respond yes they have been pregnant or gotten someone pregnant are coded as 1, those who don't know as 2*, and all others are coded as 0.				

<sup>\*</sup>Note: The sample did not include any "I don't know" (2) responses.

Table 3.2. Behavioral Outcomes Used for Secondary Impact Analyses

Outcome Name	Outcome Description	Timing of Measure Relative to Program
Knowledge about STDs	The variable is a measure of the percentage of correct responses to items regarding STD knowledge. The measure is constructed from the following three items on the survey:	Pre-Survey, immediately following program completion, six months
	<ul> <li>"The surest way to prevent pregnancy and sexually transmitted diseases, including HIV, is to avoid all forms of sexual intimacy."</li> </ul>	following program completion and one year after the completion of the program.
	<ul> <li>"The only way that women are infected with HIV is through sexual intimacy with men."</li> </ul>	
	<ul> <li>"A person with a sexually transmitted disease always has symptoms."</li> </ul>	
	The variable is constructed as a continuous variable, based on the mean scores of the above items. The values range from 0% correct to 100% correct,	
Perceived barriers to sexual health	The variable is a measure of the extent to which a person agrees or disagrees with statements regarding barriers to using birth control. The measure is constructed from the following seven items on the survey:	Pre-Survey, immediately following program completion, six months following program completion
	"In general, birth control is too much trouble to use."	and one year after the completion of the program.
	<ul> <li>"In general, birth control is too expensive to buy."</li> </ul>	
	<ul> <li>"It takes too much planning ahead of time to have birth control on hand when you're going to have sex."</li> </ul>	
	<ul> <li>"It would be too hard to get a boy to use birth control with you."</li> </ul>	
	<ul> <li>"For you, using birth control interferes with sexual enjoyment."</li> </ul>	
	<ul> <li>"It is easy for you to get birth control."*</li> </ul>	
	<ul> <li>"If you used or carry birth control, your friends might think that you were looking for sex."</li> </ul>	
	The variable is constructed as a continuous variable. Using a Likert scale from 1 to 5 (where 1 = strongly disagree and 5 = strongly agree), an average of the seven survey items is computed.	
Intention to engage in sexual intercourse	The variable is a yes/no measure (where a yes response is indicated by "yes, probably" or "yes definitely", and a no response is indicated by "no, probably not" or "no, definitely not") of whether a person intends to have sex within the next year. The measure is taken directly from the following item on the survey:	Pre-Survey, immediately following program completion, six months following program completion and one year after the
	<ul> <li>"Do you plan to have sexual intercourse in the next year, if you have the chance?"</li> </ul>	completion of the program.
	The variable is constructed as a dummy variable where respondents who respond "yes, probably" or "yes, definitely" are coded as 1, and those who respond "no, probably not" or "no, definitely not" are coded as 0.	

Outcome Name	Outcome Description	Timing of Measure Relative to Program
Intentions to use birth control	The variable is a yes/no measure (where a yes response is indicated by "yes, probably" or "yes definitely", and a no response is indicated by "no, probably not" or "no, definitely not") of whether a person plans to use birth control with the next year (should they have sex). The measure is taken directly from the following items on the survey:	Pre-Survey, immediately following program completion, six months following program completion and one year after the completion of the program.
	<ul> <li>"If you were to have sexual intercourse in the next year, do you plan to use (or have you partner use) any of these methods of birth control: Condoms, birth control pills, the shot, the patch, the ring, IUD, or implants?"</li> </ul>	
	<ul> <li>"If you have sexual intercourse in the next year, do you plan to use (or have your partner use) a condom?"</li> </ul>	
	The variable is constructed as a dummy variable where respondents who respond "yes, probably" are coded as 3, those who respond "yes, definitely" are coded as 4, those who respond "no, probably not" are coded as 2, and those who respond "no, definitely not" are coded as 1. The variable is then dichotomized, where a "yes" response on one or both is coded as 1, and a no response on both is coded as 0.	

\*Note: reverse coded (1 = strongly agree, 5 = strongly disagree)

## **Study Sample**

Overall, a total of 1,292 middle school girls were recruited during the study. Of those girls, 1,109 returned consent forms, and 998 were randomized (500 intervention, and 498 counterfactual); 111 girls did not attend the pre-session (Cohorts 3 and 4) and therefore were not randomized. As part of the randomization process, a large number of the girls were randomized in pairs (90 intervention, 122 counterfactual). To adjust for clustering, one girl from each cluster was randomly selected to be included in the analysis. As a result, 89 of the participants (for whom the evaluation team had data) were excluded from the analysis. The total number of participants used in the primary and secondary analysis samples was 803 (406 intervention, 397 counterfactual).

Approximately 72.9% (363 intervention, 365 counterfactual) of the girls completed a pre-survey, 60.2% (296 intervention, 305 counterfactual) completed an immediate post-survey, 57.1% (281 intervention, 289 counterfactual) completed a six-month post survey, and 54.2% (269 intervention, 272 counterfactual) completed a twelve-month post survey (see Table C in Appendix C).

To assess the primary outcome measures, the evaluation team used an analytic sample consisting of cases that completed both a pre-survey and a 12-month post-survey. Data for those who did not have a pre-survey were imputed to a constant (0), and a dummy variable was created (0 = non imputed score, 1 = imputed score) for each of the primary outcome measures. The final analytic sample for the primary impact analysis includes 541 cases from 34 sites pooled across all cohorts, with 269 cases in the intervention condition and 272 cases in the counterfactual condition (see Table C.1 in Appendix C).

To assess the secondary outcome measures, the evaluation team used multiple analytic samples consisting of cases who completed both a pre-test survey and an immediate post-survey, those who completed both a pre-survey and six-month follow up survey, and those who completed both a pre-survey and twelve month post-survey. For those who did not complete a pre-survey, data for the variables of interest were imputed using a constant (0), and a dummy variable for each outcome of interest was created (0 = non imputed score, 1 = imputed score). The analytic pre and immediate post sample includes 601 cases from 34 sites pooled across all cohorts (296 intervention and 305 counterfactual), The analytic pre and six-month post-survey sample includes 570 cases from 34 sites pooled across all cohorts (281 intervention, 289 counterfactual; see Tables C.2 and C.3 in Appendix C).

#### **Pre-survey Equivalence**

To ensure that the randomization process was completed successfully, and that effects seen in the post-tests were not due to biases that existed at pre-survey or that arose due to attrition, the evaluation team conducted multivariate and binary logistic regressions for the demographic variables (age, race and gender) and the primary outcomes collected during the pre-survey for each analytic sample (see Tables 4, 5 and 6). Pre-survey differences on the secondary outcomes (i.e., STD knowledge, perceived barriers to sexual health, intent to engage in sexual intercourse,

intention to use birth control) were also examined using multivariate and binary logistic regressions, with the results generally indicating non-significant differences (see Tables D.1. through D.3. in Appendix D).

A *p*-value of .05 (two-tailed) was set as the criteria for significance on the primary and secondary outcome measures at pre-survey.

The evaluation team found no pre-survey nonequivalence, among the three analytic samples.

Therefore, the team deemed it unnecessary to perform any matching procedures and proceeded with the impact analysis.

Table 4. Summary Statistics of Primary Pre-survey Measures for Girls Completing Immediate Post-survey

Pre-survey Measure	Intervention Mean or % (Standard Deviation)	Counterfactual Mean or % (Standard Deviation)	Intervention Versus Counterfactual Mean Difference (p-value of difference)	Sample Size (Intervention/ Counterfactual)
Age	11.58 (.69)	11.54 (.73)	0.04 (.57)	549 (267/282)
Gender (female)	100%	100%	.00 (1.0)	549 (267/282)
Race/ethnicity: Latina	76.5%	82.0%	06 (.14)	493 (243/250)
Race/ethnicity: Asian	3.3%	1.2%	.02 (.12)	493 (243/250)
Race/ethnicity: Black	6.2%	6.4%	00 (.92)	493 (243/250)
Race/ethnicity: White	3.7%	2.8%	.01 (.57)	493 (243/250)
Race/ethnicity: Two or more	9.1%	5.2%	.04 (.10)	493 (243/250)
Race/ethnicity: Other	1.2%	2.4%	01 (.34)	493 (243/250)
Sexual Activity Onset (% yes)	0.0%	0.004%	00 (.34)	549 (267/282)
Pregnancy Incidence (% yes)	0.0%	0.0%	.00 (1.0)	549 (267/282)

Table 5. Summary Statistics of Primary Pre-survey Measures for Girls Completing a Six-Month Post-Survey

Pre-survey Measure	Intervention Mean or % (Standard Deviation)	Counterfactual Mean or % (Standard Deviation)	Intervention Versus Counterfactual Mean Difference (p-value of difference)	Sample Size (Intervention/ Counterfactual)
Age	11.53 (.71)	11.57 (.72)	04 (.51)	549 (267/282)
Gender (female)	100%	100%	.00 (1.0)	549 (267/282)
Race/ethnicity: Latina	75.6%	82.6%	07 (.06)	475 (234/241)
Race/ethnicity: Asian	3.4%	1.2%	.02 (.13)	475 (234/241)
Race/ethnicity: Black	4.7%	3.3%	.01 (.45)	475 (234/241)
Race/ethnicity: White	2.6%	2.9%	00 (.82)	475 (234/241)
Race/ethnicity: Two or more	11.1%	6.6%	.04 (.09)	475 (234/241)
Race/ethnicity: Other	2.6%	3.3%	01 (.63)	475 (234/241)
Sexual Activity Onset (% yes)	0.0%	0.0%	.00 (1.0)	523 (254/269)
Pregnancy Incidence (% yes)	0.0%	0.0%	.00 (1.0)	523 (267/282)

Table 6. Summary Statistics of Primary Pre-survey Measures for Girls Completing a Twelve-Month Post-Survey

Pre-survey Measure	Intervention Mean or % (Standard Deviation)	Counterfactual Mean or % (Standard Deviation)	Intervention Versus Counterfactual Mean Difference (p-value of difference)	Sample Size (Intervention/ Counterfactual)
Age	11.63 (.727)	11.62 (.709)	.02 (.80)	485 (237/248)
Gender (female)	100%	100%	.00 (1.0)	485 (237/248)
Race/ethnicity: Latina	76.0%	81.3%	05 (.18)	436 (217/219)
Race/ethnicity: Asian	3.2%	1.4%	.02 (.21)	436 (217/219)
Race/ethnicity: Black	5.5%	5.0%	.01 (.81)	436 (217/219)
Race/ethnicity: White	3.2%	3.2%	.00 (.99)	436 (217/219)
Race/ethnicity: Two or more	10.6%	5.9%	.05 (.08)	436 (217/219)
Race/ethnicity: Other	1.4%	3.2%	02 (.22)	436 (217/219)
Sexual Activity Onset (% yes)	0.4%	0.8%	00 (.60)	485 (237/248)
Pregnancy Incidence (% yes)	0.0%	0.0%	.00 (1.0)	485 (237/248)

#### Methods

#### Impact Evaluation

The external evaluation team used an Analyses of Covariance (ANCOVA) to address primary and secondary questions. The team included girls' ages, races and pre-test scores (in addition to a dummy coded indicator of pre-survey imputation) as covariates. Sample weights were not used in these analyses since randomization occurred at the individual student level, and a selection into the study was not weighted on any characteristics. Additionally, the team adjusted significance criteria for the primary and secondary outcome measures using a Bonferroni correction, where the alpha level is divided by the number of hypotheses tests conducted. Given the number of outcomes addressed in the primary (2) research questions, the alpha was set at .025. There was no alpha adjustment for the 16 secondary research questions. An additional summary about the data management, missing data, and model specification procedures can be reviewed in Appendix E.

The evaluation team defined its analytic sample to address the primary research questions as girls who completed both a pre-survey and 12-month post-survey pooled across all cohorts. To address the secondary research questions, the team defined its analytic sample as girls who completed a 1) pre-survey and immediate post-survey, 2) a pre-survey and six-month post-survey, and 3) a pre-survey and

twelve-month post-survey, pooled across all cohorts. The program impact on sexual activity onset and pregnancy incidence was determined using an ANCOVA to compare the proportion of girls in WPWP who had previously engaged in sexual intercourse, and/or have been pregnant, to girls in EESS, at one year after completing the program as primary research questions and immediately following the program, six months after program completion as secondary research questions.

The team conducted an ANCOVA to assess the impact of the WPWP program on sexual health knowledge (knowledge of STDs), and perceived barriers to sexual health, intentions to engage in sexual intercourse, and intentions to use birth control – secondary research questions. Analyses were conducted immediately following the program, six months after program completion, and one year after program completion. The team conducted additional sub-analyses to test the measures' sensitivity.

# Implementation Evaluation

The potential impact of program implementation on primary and secondary outcomes was measured according to program adherence and program quality.

**Program adherence.** The evaluation team determined program adherence using three different methods: 1) program dosage, 2) program content and 3) program delivery. The team calculated dosage using the number of sessions offered at each site (per cohort), and the number of sessions attended (per cohort). Having previously decided that a minimum of 8 sessions attended was necessary to achieve program fidelity, a cutoff value based around 8 sessions attended was created. Attendance data were dummy coded such that participants who attended 8 or more sessions were marked 1 and participants who attended seven sessions or less were marked 0. Percentages were calculated looking at how many students from each cohort attended an adequate amount of sessions.

Program content was calculated using the total number of topics offered per site, in addition to a comparison of actual topics covered (see Appendix H for list of session names and topics covered).

Program delivery was calculated using the number of program specialists and their replacements<sup>3</sup> to produce an average number of substitute days per classroom, per cohort. Additionally, the number of Program Specialists who received training was used to calculate the percentage of trained staff.

**Program Quality.** Program quality, using observations, was determined by gauging the quality of staff-participant interactions and the quality of engagement with the program. Data collectors used a 5-point Likert scale (where ratings approaching 1 are deemed low and those approaching 5 are considered high) to assess quality components in the classrooms. They focused on one WPWP session where they would not distract learning while observing. The quality of staff-participant interactions was calculated using the percentage of observation guide items (e.g., rapport and communication with participants, effectively addressed questions/concerns, created a welcoming climate that honors and respects differences) for which data collectors rated as being *high*. The quality of engagement with the program was calculated using the percentage of observation guide items (e.g., to what extent did the participants appear to understand the material, how actively did the group members participate in discussions and activities) for which data collectors rated as being *high*.

# **Study Findings**

#### **Impact Study Findings**

#### **Primary Outcome Measures**

Table 7 shows the estimated effect of WPWP on the primary outcome measures one year after program completion. The analyses indicated that there is no evidence to suggest WPWP has any long-term impacts on sexual activity onset or pregnancy incidence.

23

<sup>&</sup>lt;sup>3</sup> A *replacement staff* is a Program Specialist who assumes leadership over a classroom to which she was not originally assigned for the semester's remainder.

Table 7. Post-Intervention Estimated Effects Using Data from Twelve-Month Post-Survey to Address the Primary Research Questions

Outcome Measure	Intervention Mean or % (Standard Deviation)	Counterfactual Mean or % (Standard Deviation)	Intervention Compared to Counterfactual Mean Difference (p-value of Difference)	Sample Size (Intervention/ Counterfactual)
Sexual Activity Onset (% yes)	0.0%	.01%	01 (.19)	494 (241/253)
Pregnancy Incidence (% yes)	0.0%	0.0%	.00 (1.0)	494 (241/253)

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

Note: Analysis controlled for the following variables: Age, Race, Pre-test scores, and the dummy coded imputed Pre-

Survey.

# **Secondary Outcome Measures**

Tables 8, 9, and 10 show the estimated effect of the secondary outcome measures, immediately following the program, six-months following program completion, and one year following program completion. Similar to the results of the primary research questions, there is no evidence to suggest that WPWP has any short-term impact on sexual activity onset or pregnancy incidence. Additionally, there is no evidence to suggest that WPWP has any short-term or long-term impacts on the girls' perception of barriers to sexual health, intentions to engage in sexual intercourse, or intentions to use birth control. However, the evidence does suggest that WPWP has both short-term and long-term impacts on girls' knowledge of STDs. The significant difference between groups on STD knowledge at the immediate post-survey (15.77 point difference; p < .001), six-month post-survey (7.56 point difference; p = .001), and twelve-month post-survey (10.68 point difference; p < .001), indicates that girls within WPWP select a significantly higher proportion of correct responses (overall), when compared to girls in EESS, up to one year after the program.

Table 8. Post-Intervention Estimated Effects Using Data from Immediate Post-Survey to Address the Secondary Research Questions

Outcome Measure	Intervention Mean or % (Standard Deviation)	Counterfactual Mean or % (Standard Deviation)	Intervention Compared to Counterfactual Mean Difference (p-value of Difference)	Sample Size (Intervention/ Counterfactual)
Sexual Activity Onset (% yes)	0.0%	0.0%	00 (.34)	561 (279/282)
Pregnancy Incidence (% yes)	0.0%	0.0%	.00 (1.0)	561 (279/282)
Knowledge about STDs	32.81 (28.50)	17.05 (23.19)	15.76 (.00)*	516 (258/258)
Perceived Barriers to Sexual Health	2.83 (.58)	2.90 (.49)	06 (.14)	435 (220/215)
Intention to Engage in Sexual Intercourse (% yes)	.02%	.02%	01 (.73)	506 (254/252)
Intention to use Birth Control (% yes)	86%	83%	.03 (.31)	433 (216/217)

Source: Follow-up surveys administered immediately after the program.

Note: Analysis controlled for the following variables: Age, Race, Pre-test scores, and the dummy coded imputed Pre-Survey. \* = statistically significant at p < .01.

Table 9. Post-Intervention Estimated Effects Using Data from Six-Month Post-Survey to Address the Secondary Research Questions

Outcome Measure	Intervention Mean or % (Standard Deviation)	Counterfactual Mean or % (Standard Deviation)	Intervention Compared to Counterfactual Mean Difference (p-value of Difference)	Sample Size (Intervention/ Counterfactual)
Sexual Activity Onset (% yes)	0.0%	0.0%	.00 (.91)	540 (267/273)
Pregnancy Incidence (% yes)	0.0%	0.0%	.00 (1.0)	540 (267/273)
Knowledge about STDs	32.13 (30.01)	24.58 (28.01)	7.55 (.001)*	506 (251/255)
Perceived Barriers to Sexual Health	2.83 (.556)	2.87 (.53)	04 (.49)	452 (226/226)
Intention to Engage in Sexual Intercourse (% yes)	.01%	.01%	.00 (.45)	503 (253/250)
Intention to use Birth Control (% yes)	86%	85%	.01 (.48)	432 (214/218)

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

Note: Analysis controlled for the following variables: Age, Race, Pre-test scores, and the dummy coded imputed Pre Survey. \* = statistically significant at p < .01.

Table 10. Post-Intervention Estimated Effects Using Data from Twelve-Month Post-Survey to Address the Secondary Research Questions

Outcome Measure	Intervention Mean or % (Standard Deviation)	Counterfactual Mean or % (Standard Deviation)	Intervention Compared to Counterfactual Mean Difference (p-value of Difference)	Sample Size (Intervention/ Counterfactual)
Knowledge about STDs	35.89 (30.39)	25.21 (26.49)	10.68 (.000)*	456 (222/234)
Perceived Barriers to Sexual Health	2.84 (.54)	2.85 (.49)	01 (.78)	414 (203/211)
Intention to Engage in Sexual Intercourse (% yes)	.04%	.02%	.02 (.08)	468 (232/236)
Intention to use Birth Control (% yes)	90%	85%	.05 (.08)	399 (192/207)

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

Note: Analysis controlled for the following variables: Age, Race, Pre-test scores, and the dummy coded imputed Pre-Survey. \* = statistically significant at p < .001.

## **Sensitivity Analysis**

To investigate the sensitivity of the conducted analyses in finding significant results, the evaluation team conducted two different sub-analyses; (1) an analysis of the primary and secondary research outcomes, having removed the covariates (e.g., age, race, pre-test scores); and (2) an analysis of the primary and secondary research outcomes using a sub-sample of girls who attended 7 or more program sessions (N = 617; 313 intervention, 304 counterfactual).

The first analysis involved performing chi-square and t-test analyses on the primary and secondary research outcomes. The results indicated no differences when compared to the reported findings (see Tables F.1-F.4 in Appendix F for full results).

To conduct the second analysis a dummy variable was created, where all girls who attended 7 or more of the program sessions were coded as 1, and the girls who attended 6 or fewer program sessions (thus having missed 4 or more sessions) were coded as 0. An ANCOVA (with age, race, gender, pre-test scores, and a dummy coded imputed pre-test variable as covariates) was used to assess the impact of the WPWP program on both primary and secondary research questions. The results indicated no differences (compared to previous findings) among the primary or secondary research questions (see Tables F.5 – F.8 in Appendix F). The results of both sub-analyses indicate a sufficient sensitivity within the original analyses in finding significant differences.

# **Implementation Study Findings**

#### **Implementation Quality**

#### Program Adherence

To obtain information on program adherence, the evaluation team used attendance to examine session attendance and number of sessions offered (dosage), fidelity logs to explore the type and number of topics covered (content), and teaching specialist training information (delivery).

**Dosage.** Cohort 1 data indicated that 70.5% (N = 44) of girls in the intervention group, and 70.3% (N = 37) of girls in the counterfactual group completed 8 or more sessions of the program. In Cohort 2, 66.7% (N = 108) of the girls in the intervention group and 73.1% (N = 104) of the girls in the counterfactual group completed 8 or more sessions. In Cohort 3, 77.1% (N = 144) of the girls in the intervention group and 76% (N = 150) of the girls in the counterfactual group completed 8 or more sessions. Finally, in Cohort 4, 76.9% (N = 104) of the girls in the intervention group and 66.7% (N = 105) of the girls in the counterfactual group completed more than 8 sessions. Overall, 73.5% (N = 400) of girls in the intervention group, and 72.2% (N = 396) of girls in the counterfactual group completed 8 or more sessions. Refer to Appendix G for a dosage summary.

*Content.* With 11 WPWP sessions and 10 EESS sessions implemented during each cohort, Program Specialists taught 714 sessions (374 intervention and 340 counterfactual) over the grant period.

**Delivery.** The data regarding the number of trained program specialists indicated that 100% of the specialists received training (7 intervention, 7 counterfactual), of which, all were responsible for program delivery. Of the total number of sessions offered (11 intervention, 10 counterfactual), eight sessions were delivered by a short-term (1-2 days), trained replacement staff, and 12 sessions were delivered by a replacement staff over the study's course. A *replacement staff* is a Program Specialist who assumes leadership over a classroom to which she was not originally assigned for the semester's remainder. Additionally, ongoing support was offered throughout the duration of the program. The Program Manager led weekly staff meetings with Program Specialists. Approximately 15-17 staff meetings took place for each cohort. The manager conducted six site visits during Cohort 1, 10 visits during Cohort 2, 6 visits during Cohort 3 and 6 visits during Cohort 4.

## **Program Quality**

Among the 28 unduplicated observations conducted, data collectors indicated that 79% of the observed staff-participant interactions were considered to be of *higher quality* (a rating of 4 or 5, out of 5) and only 21% of the observed staff-participant interactions had rating indicating very low engagement (2 of 5) or moderate engagement (3 of 5). When looking at youth engagement, it was reported that 87% of the girls were considered *highly engaged* (a rating of 4 or 5, out of 5) and 13% being rated as having very low engagement (2 of 5) or moderate engagement (3 of 5).

Per participants' feedback, the most memorable activity during the sessions was role-playing. Through these skits, girls reported that they learned how to leave situations without being rude but still being assertive. The second most memorable activity involved learning about different parts of the female anatomy. In addition, girls also disclosed that they learned about 1) the risk of pregnancy, STDs and HIV and 2) the Bill of Rights for Women. When girls were asked about the parent/daughter workshop, they reported that they learned what their parents knew about sex and topics related to sex. Some girls admitted that it was weird or awkward discussing sexual topics with parents.

# **Sexual Intentions**

During the focus groups, data collectors provided girls with the following scenario to understand what advice that they would give to a friend in a potentially risky situation in which a *hottie* (a physically and/or romantically attractive peer) was invited over when her parents were not home. Many girls expressed their reservations about their friend inviting someone over without parental supervision. They explained that they would tell their friend to think about her intentions in inviting the hottie over and to consider the possible consequences of the situation. Girls were most concerned with the risk of their friend engaging in sexual intercourse because it could lead to pregnancy.

#### **Bill of Rights for Women**

The focus group guide included a few questions that tested girls' retention and application of the material that was presented to them in the WPWP sessions program, specifically the Bill of Rights for Women. A handful of girls in the focus groups reported that they used the Bill of Rights for Women to say no to situations. The situations that girls shared included refusing requests from family members to do chores around the house, saying no to cheating on an exam, rejecting an offer to try hookah, and refusing to hang out with a friend when she invited a boy over to watch a movie.

#### **Consequences of Sex and Pregnancy**

Data collectors asked girls what would change in someone's life – and in their own lives – if they learned they were pregnant. Girls reported many consequences of sex and pregnancy, the most popular being the inability for a pregnant teen to complete her education, increased responsibility, the need for financial support, and the need to find a job to support the baby. Girls also noted that their parents would treat them differently; specifically that their parents and friends would be upset and shocked.

Access to Sexual Health Information and Resources. Girls explained that young girls could find information about sexual health and access to birth control by asking their doctor, asking their parents, asking a GIGLA specialist, asking a trusted adult, and asking a Planned Parenthood staff.

#### **Roles of Religion and Culture**

In the targeted focus groups, data collectors asked one question about how the girls' religion and culture have shaped their knowledge of sex and sexual health. Most of the girls who answered this question discussed how their religion taught them to abstain from sex before marriage and to remain a virgin. Girls also discussed the social norms in their home and mentioned that they did not talk about sex or sexual health with their family members. Some attributed this to their religion while others did not have a reason.

#### **Conclusions**

Key outcomes of the implementation evaluation suggest a high fidelity of program implementation. On average, girls in both intervention and counterfactual conditions attended a majority (8 or more) of their program sessions. Session observations revealed a high level of rapport among teachers and students, in addition to high levels of understanding of the material and participation. Additionally, all of program specialists received training, and most were present for all sessions.

Primary impact analyses indicated that WPWP has no long-term impact on sexual activity onset or pregnancy incidence when compared to EESS. The secondary analyses revealed similar results –WPWP has no impact (short or long-term) on girls' perception to sexual health, intention to engage in sexual intercourse, or intention to use birth control, when compared to EESS. The analyses did indicate that when compared to EESS participants, those completing WPWP have greater STD knowledge both immediately following the program end, six months following the program, and one year after the program.

In looking at possible explanations for these differences, or lack thereof, age should be considered. Non-significant results on primary outcomes (sexual activity and pregnancy) are almost certainly expected among middle school aged girls. The focus would then fall to sexual health knowledge and intentions. WPWP participants scored higher on STD knowledge, however there were no differences in their intentions to engage in sex or use birth control, suggesting that knowledge comes with exposure. It could be argued that as girls grow older and engage in sexual discussions or encounters, they will obtain that knowledge. When controlling for age (accounting for the variance between groups that is explained by age), the differences in STD knowledge was still significantly different, meaning that age was unimportant. To note, the

change in age (year of growth) at this level was minimal. It might help to look at the possible effect of age in longer studies.

Another issue to consider is the implementation of the RCT design, where samples were not stratified geographically or by school. Limitations arose regarding the evaluation team's ability to control the information available to students. While the evaluation team's intent was to assign any participants (individuals or clusters) to one condition or the other to reduce exposure to the material through friends, both intervention and counterfactual groups existed in the same schools and so diffusion of course material may have occurred through simple conversation between peers outside of the program setting. This diffusion has clear implications on the design's power to detect differences between the counterfactual and intervention groups.

#### References

- Advocates for Youth (2009). *Comprehensive sex education: Research and results*. Retrieved on October 29, 2016 from http://www.advocatesforyouth.org/storage/advfy/documents/fscse.pdf
- California Department of Public Health (2009). *California Adolescent Health 2009*. Retrieved on February 20, 2016 from https://www.cdph.ca.gov/pubsforms/Pubs/OWH-AdolHealthReport09.pdf
- California Department of Public Health (2013). "California Teen Birth Rates, 1991-2011." {Slideshow}. Retrieved on February 20, 2016 from https://www.cdph.ca.gov/programs/mcah/Documents/MO-MCAH-2011TBR-DataSlides.pdf
- Carmichael, Paul H. (1997). "Who Receives Federal Title I Assistance?: Examination of Program Funding by School Poverty Rate in New York State." *Educational Evaluation and Policy Analysis*, 19 (4): 354-359.
- Centers for Disease Control and Prevention (2007). *Youth Risk Behavior Surveillance United States*, 2007. Retrieved on October 29, 2015 from <a href="http://www.cdc.gov/mmwr/preview/mmwrhtml/ss5704a1.htm">http://www.cdc.gov/mmwr/preview/mmwrhtml/ss5704a1.htm</a>
- Centers for Disease Control and Prevention (2009). *Sexually Transmitted Disease Surveillance*, 2008. Retrieved on October 29, 2015 from http://www.cdc.gov/std/stats08/surv2008-complete.pdf.
- Ho, D., Imai, K., King, G., Stuart, E. (2011). MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software*, 42 (8), 1-28. https://www.cdph.ca.gov/pubsforms/Pubs/OWH-AdolHealthReport09.pdf
- Eaton, D., Kann, L., Kinchen, S., Shanklin, S., Ross, J., Hawkins, J., Harris, W., Lowry, R., McManus, T., Chyen, D., Lim, C., Brener, N., and Wechsler, H. (2007). Youth Risk Behavior Surveillance—United States 2007. Morbidity and Mortality Weekly Report (June, 2008); 1-131.
- "Girls Inc. Preventing Adolescent Pregnancy." Girls Inc. Retrieved on February 20, 2016 from http://www.girlsinc.org/resources/programs/girls-inc-preventing-adolescent-pregnancy.html.
- Public Health Institute (2008, May). *No time for complacency: Teen births in California*. Retrieved on October 29, 2015 from http://teenbirths.phi.org/2008TeenBirthsReport.pdf.
- "TPP Research and Demonstration Programs." Department of Health and Human Services. Retrieved on February 20, 2016 from http://www.hhs.gov/ash/oah/oah-initiatives/teen pregnancy/about/research-demonstration.html.
- University of California San Francisco (n.d.). A Question of Hope: Reducing Latina Teen Childbearing in California.

#### **APPENDIX A**

# **DATA COLLECTION EFFORTS**

Table i. California Teen Birth Rates Per 1,000 Girls Ages 15 to 19 by Age Group (California Department of Health, 2013)

	Year			
Age	2009	2010	2011	
15-17	19.2	16.4	14.8	
15-19	35.4	31.5	28.0	
18-19	59.2	53.1	46.7	

Table ii. California Teen Birth Rates Among Girls 15-19 by Race or Ethnicity (California Department of Health, 2013)

	Year			
Race or Ethnicity	2000	2005	2011	
African American	81,973	90,880	87,047	
Asian	139,798	145,237	148,390	
Hispanic	477,270	552,682	662,968	
Pacific Islander	4,829	5,402	5,421	
White	442,794	455,967	409,211	

Table A.1. Data collection efforts used in the impact analysis of Will Power/Won't Power and timing

	Randomization	Pre-survey	End of Program Post-survey	6 Months Post-survey	12 Months Post- survey
Cohort 1	10/22/12-	10/30/12-	1/12/13-	8/13/13-	1/29/14-
	11/12/12	11/16/12	3/8/13	09/15/13	2/29/14
Cohort 2a	2/4/13-	2/5/13-	5/28/13-	11/18/13-	5/5/14-
	3/15/13	3/15/13	6/13/13	12/2/13	5/16/14
Cohort 2b	3/4/13-	3/6/13-	5/28/13-	12/4/13-	5/19/14-
	3/21/13	3/21/13	6/6/13	12/18/13	5/30-14
Cohort 3	9/10/13-	9/10/13-	12/3/13-	8/11/14-	01/12/15-
	10/16/13	10/16/13	1/22/13	9/12/14	02/6/15
Cohort 4	2/4/14-	2/4/14-	4/30/13-	11/3/14-	04/14/15-
	3/17/14	3/17/14	6/7/13	11/21/14	05/22/15

# **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	
	Ac	dherence		
(1) How many and how often were sessions offered	All sessions are recorded in attendance sheets and fidelity monitoring logs.	Attendance sheets are completed during each session and fidelity monitoring logs are completed after each session.	Program Specialists	
(2) What and how much was received	Attendance sheets capture the number of sessions each girl attends and the length of each session.	Participants log their time of arrival and departure to indicate the amount of time spent at each session.	Program Specialists	
(3) What content was delivered to youth	Fidelity monitoring logs capture the unique content topics for each session.	Fidelity monitoring logs are recorded after every session.      A fidelity monitoring log is completed after observation of	Program     Specialists  2. AES Data	
		Session 6 - The Case for Abstinence at each site.	Collectors (external evaluation team)	
(4) Who delivered material to youth	Resumes and applications of JobScore (online human resource [HR] site).	Resumes collected prior to employment interview.	1.HR & Program Manager	
	Training agendas and sign-in sheet.	Agendas and sign-in sheets collected at each training session. (All trainings are required.)	Program Manager     HR	
	Position requirements posted on job postings.	3. Job postings on an as- needed basis.		
		ality		
Quality of staff- participant interactions	Data Collectors observe classrooms to assess implementation quality and curriculum fidelity using a modified, quantitative observation guide with	Data Collectors observe one specific WPWP session – Session 6: The Case for Abstinence – in each WPWP classroom	AES Data Collectors	
	qualitative fields for elaboration.	All girls in attendance are observed.		

Implementation Element  Quality of youth engagement with program	Types of data used to assess whether the element of the intervention was implemented as intended  Combination of information collected from classroom observations and focus groups (youth self-report) to accurately report youth's program engagement.	Frequency/sampling of data collection  Observations: Data Collectors observe one specific session (Session 6: The Case for Abstinence) in each WPWP classroom All girls in attendance during the session are observed.  Focus groups: At the close of each cohort, the external evaluation team conducts 4-6 focus groups with WPWP participants at purposefully selected sites  Program Specialists alert girls to the opportunity to join one site-specific focus group one week following Post-survey 1. Any WPWP participant may volunteer to return for the focus group. Girls are offered a gift card as a "thank you" for participation.	Party responsible for data collection AES Data Collectors	
Experiences of counterfactual condition			AES Data Collectors	
Other TPP		ntext NA	NA	
programming available or offered to study participants (both T and C)	All 7th-graders receive state- mandated health course.			
External events affecting implementation	Fidelity Monitoring logs     track site specific issues     Staff meeting notes record     issues within the school     districts     Calendars and     announcements on school     websites	1. Ad hoc 2. Weekly 3. Weekly	Program     Specialist     Program     Assistant     Program     Specialist	
Substantial unplanned adaptation(s)	Workplan adaptation documented in Scope Change Proposal and Revised Year 4 Workplan	Ad hoc	Program Management Team	

# APPENDIX C

# STUDY SAMPLE

Table C. Middle School Analytic Samples for Primary and Secondary Outcome Measures

	Time Period	Total sample size	Intervention sample size	Comparison sample size	Total response rate	Intervention response rate	Counterfactual response rate
Number of Youth							
Assigned to condition		998	500	498			
2. Contributed a Pre-Survey		728	363	365	72.9%	72.6%	73.2%
3. Contributed a follow-up survey	Immediately post- programming	601	296	305	60.2%	59.2%	61.2%
4. Contributed a follow-up survey	6 months post- programming	570	281	289	57.1%	56.2%	58.0%
5. Contributed a follow-up survey	12 months post- programming	541	269	272	54.2%	53.8%	54.6%

Table C.1. Middle School Cases with Pre and Twelve-Month Post-Survey

Cohort	# of Sites	# of Counterfactual Cases	# of Intervention Cases	Total # of Cases
1	4	32	37	69
2	8	86	83	169
3	12	83	70	153
4	10	71	79	150
Total	34	272	269	541

Table C.2. Middle School Cases with Pre and Immediate Post-Survey

Cohort	# of Sites	# of Counterfactual Cases	# of Intervention Cases	Total # of Cases
1	4	33	35	68
2	8	79	75	154
3	12	118	106	224
4	10	75	80	155
Total	34	305	296	601

Table C.3. Middle School Cases with Pre and Six-Month Post-Survey

Cohort	# of Sites	# of Counterfactual Cases	# of Intervention Cases	Total # of Cases
1	4	25	22	47
2	8	85	74	159
3	12	94	96	190
4	10	85	89	174
Total	34	289	281	570

## **APPENDIX D**

## PRE-SURVEY EQUIVALENCIES OF SECONDARY OUTCOME MEASURES

Table D.1. Summary Statistics of Secondary Pre-Survey Measures for Girls Completing Immediate Post-Survey

Pre-Survey Measure	Intervention Mean or % (Standard Deviation)	Comparison Mean or % (Standard Deviation)	Intervention Versus Comparison Mean Difference (p-value of difference)	Sample Size (Intervention/ Comparison)
Knowledge about STDs	19.58 (24.63)	18.70 (22.46)	.87 (.67)	514 (252/262)
Perceived Barriers to Sexual Health	2.94 (.552)	2.96 (.56)	02 (.73)	456 (218/238)
Intention to Engage in Sexual Intercourse (% Yes)	0.4%	3.4%	03 (.04)	515 (251/264)
Intention to use Birth Control (% Yes)	73.4%	72.5%	.01 (.84)	447 (214/23)

Table D.2. Summary Statistics of Secondary Pre-Survey Measures for Girls Completing Six-Month Post-Survey

Pre-Survey Measure	Intervention Mean or % (Standard Deviation)	Comparison Mean or % (Standard Deviation)	Intervention Versus Comparison Mean Difference (p-value of difference)	Sample Size (Intervention/ Comparison)
Knowledge about STDs	18.74 (24.25)	18.99 (23.24)	26 (.90)	493 (242/251)
Perceived Barriers to Sexual Health	2.89 (.55)	2.95 (.54)	07 (.21)	448 (217/231)
Intention to Engage in Sexual Intercourse (% Yes)	0.4%	2.4%	02 (.10)	496 (242/254)
Intention to use Birth Control (% Yes)	71.7%	75.4%	03 (.37)	440 (212/228)

Table D.3. Summary Statistics of Secondary Pre-Survey Measures for Girls Completing Twelve-Month Post-Survey

Pre-Survey Measure	Intervention Mean or % (Standard Deviation)	Comparison Mean or % (Standard Deviation)	Intervention Versus Comparison Mean Difference (p-value of difference)	Sample Size (Intervention/ Comparison)
Knowledge about STDs	18.81 (24.33)	19.37 (23.20)	56 (.80)	459 (225/234)
Perceived Barriers to Sexual Health	2.90 (.56)	2.96 (.55)	06 (.25)	411 (200/211)
Intention to Engage in Sexual Intercourse (% Yes)	0.9%	3.4%	03 (.08)	464 (229/235)
Intention to use Birth Control (% Yes)	73.2%	74.4%	01 (.78)	405 (194/211)

#### APPENDIX E

#### DATA MANAGEMENT AND MISSING DATA PROCEDURES

SPSS software was used to analyze the data. Specifically, all survey data were entered into SPSS files, merged with past data, cleaned, and analyzed. To clean the data, inconsistencies within and between surveys were identified, reported, and treated as missing data. Specifically, if a case indicated at pre-survey that she has had sex, yet indicated at follow-up that she had not, the follow-up response was treated as missing. Alternatively, if a case indicated that they have had sex in the last three months, but did not indicate a number, both variables were treated as missing. This strategy was applied consistently to both the intervention and counterfactual groups on all three post-tests.

For each outcome, analyses determined whether there were any systematic differences on pre-survey characteristics between the cases that provided a response for the particular outcome and those who did not. Specifically, a dummy variable was created to indicate whether a given outcome variable was missing for each case. Cases missing and not missing that outcome variable were examined for equivalence on all other outcome variables and demographic characteristics (age, ethnicity). This process was repeated for all outcome variables. The missing data analysis was written in SPSS syntax and thus can be applied consistently to all outcome variables for each report. No differences were found; therefore, the analyses of the outcome data proceeded as planned.

#### MODEL SPECIFICATION

To determine the impact of the program on Sexual Activity Onset, an Analysis of Covariance (ANCOVA; controlling for age, race, pre-survey scores, and a dummy coded indication of pre-survey imputation) was conducted to compare the proportion of girls in WPWP who have had sex to girls in EESS, at one year after completing the program.

To determine the impact of the program on Pregnancy Incidence an ANCOVA (controlling for age, race, pre-survey scores, and a dummy coded indication of pre-survey imputation) was conducted to compare the proportion of girls in WPWP who have been pregnant to girls in EESS, at one year after completing the program. An ANCOVA (controlling for age, race, pre-survey scores, and a dummy coded indication of pre-survey imputation) was conducted to assess the impact of the WPWP program on sexual health knowledge and perceived barriers to sexual health, one year after program completion. Analyses compared the average STD knowledge and average perceived barriers to sexual health of girls in WPWP to those in EESS at one year after program completion.

An ANCOVA (controlling for age, race, pre-survey scores, and a dummy coded indication of pre-survey imputation) was conducted to assess the impact of the WPWP program on intentions to engage in sexual intercourse, and intentions to use sexual protection. Analyses compared the number of girls who report that they intend to have sex in the next year in the WPWP program to those in the EESS program at one year after program completion. Analyses also compared the number of girls who report that they intend to use sexual protection in the next year in the WPWP program to those in the EESS program at one year after program completion.

## 1. Software

SPSS software was used to analyze the data. Specifically, all survey data were entered into SPSS files, merged with past data, cleaned, and analyzed. SPSS syntax was used to ensure that consistent cleaning, scoring, and analysis procedures were used at all primary and secondary samples.

## 2. Criteria for statistical significance $(p \langle .05)$

As described in the body of the report, the evaluation team determined the criteria for significance, by dividing .05 by the number of hypotheses tests conducted. Based on the selected primary (2) the team anticipated this would yield a critical *p* of .025. No adjustments were made for the 16 secondary outcomes. Two-tailed tests were used to test hypotheses.

## 3. Clustering

A small number of the girls were randomized in pairs. As a clustering adjustment, the evaluation team randomly selected one girl from each cluster to be included in the analysis.

## **APPENDIX F**

## **SENSITIVITY ANALYSES**

Sensitivity Analyses - Primary and Secondary Outcomes with Covariates Removed

Table F.1. Post-intervention estimated effects using data from twelve-month post-survey to address the primary research questions

Outcome measure	Intervention mean or % (standard deviation)	Counterfactual mean or % (standard deviation)	Intervention compared to counterfactual mean difference (p-value of difference)	Sample Size (Intervention/ Counterfactual)
Sexual Activity Onset (% yes)	0.0%	1.1%	01 (.08)	541 (269/272)
Incidence of Pregnancy (% yes)	0.0%	0.0%	.00 (1.0)	541 (269/272)

Source:

Follow-up surveys administered 12 months after the program.

Table F.2. Post-intervention estimated effects using data from twelve-month post-survey to address the secondary research questions

Outcome measure	Intervention mean or % (standard deviation)	Counterfactual mean or % (standard deviation)	Intervention compared with counterfactual Mean difference (p-value of difference)	Sample Size (Intervention/ Counterfactual)
Knowledge about STDs	35.39 (30.15)	24.49 (25.93)	10.9 (.00)*	519 (259/260)
Perceived Barriers to Sexual Health	2.83 (.52)	2.85 (.48)	02 (.78)	505 (253/252)
Intention to Engage in Sexual Intercourse (% yes)	3.5%	1.6%	.02 (.16)	515 (258/257)
Intention to use Birth Control (% yes)	89.0%	82.7%	.06 (.05)	488 (245/243)

Source:

Follow-up surveys administered 12 months after the program

Note:

\* = statistically significant at  $p \langle .01$ .

Table F.3. Post-intervention estimated effects using data from immediate post-survey to address the Secondary research questions

Outcome measure	Intervention mean or % (standard deviation)	Counterfactual mean or % (standard deviation)	Intervention compared to counterfactual mean difference (p-value of difference)	Sample Size (Intervention/ Counterfactual)
Sexual Activity Onset (% yes)	0.0%	0.3%	-0.3 (.32)	601 (296/305)
Incidence of Pregnancy (% yes)	0.0%	0.0%	.00 (1.0)	601 (296/305)
Knowledge about STDs	32.18 (28.30)	16.90 (23.11)	15.28 (.00)*	577 (289/288)
Perceived Barriers to Sexual Health	2.83 (.56)	2.90 (.52)	06 (.15)	520 (260/260)
Intention to Engage in Sexual Intercourse (% yes)	1.8%	1.8%	0.00 (.99)	566 (284/282)
Intention to use Birth Control (% yes)	84.8%	80.5%	.04 (.19)	526 (270/256)

Source: Follow-up surveys administered immediately following the program.

Note: \* = statistically significant at  $p \langle .01$ .

Table F.4. Post-intervention estimated effects using data from six-month post-survey to address the Secondary research questions

Outcome measure	Intervention mean or % (standard deviation)	Counterfactual mean or % (standard deviation)	Intervention compared to counterfactual mean difference (p-value of difference)	Sample Size (Intervention/ Counterfactual)
Sexual Activity Onset (% yes)	0.4%	0.3%	0.1 (.98)	570 (281/289)
Incidence of Pregnancy (% yes)	0.0%	0.0%	.00 (1.0)	570 (281/289)
Knowledge about STDs	31.64 (30.23)	24.26 (27.65)	7.38 (.003)*	558 (275/283)
Perceived Barriers to Sexual Health	2.83 (.55)	2.88 (.53)	05 (.36)	530 (267/263)
Intention to Engage in Sexual Intercourse (% yes)	1.5%	0.7%	0.14 (.41)	544 (273/271)
Intention to use Birth Control (% yes)	85.7%	84.7%	.01 (.76)	506 (251/255)

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

Note: \* = statistically significant at  $p \langle .01.$ 

## Sensitivity Analyses – Primary and Secondary Outcomes Using High Attendance Sample

Table F.5. Post-intervention estimated effects using data from twelve-month post-survey (high attendance) to address the primary research questions

Outcome measure	Intervention mean or % (standard deviation)	Counterfactual mean or % (standard deviation)	Intervention compared to counterfactual mean difference (p-value of difference)	Sample Size (Intervention/ Counterfactual)
Sexual Activity Onset (% yes)	0.0%	0.0%	.00 (1.0)	448 (221/227)
Incidence of Pregnancy (% yes)	0.0%	0.0%	.00 (1.0)	448 (221/227)

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

Note: Analysis controlled for the following variables: Age, Race, Pre-test scores, and the dummy coded imputed Pre-

Survey. \* = statistically significant at  $p \langle .001$ .

Table F.6. Post-intervention estimated effects using data from twelve-month post-survey (high attendance) to address the secondary research questions

Outcome measure	Intervention mean or % (standard deviation)	Counterfactual mean or % (standard deviation)	Intervention compared with counterfactual Mean difference (p-value of difference)	Sample Size (Intervention/ Counterfactual)
Knowledge about STDs	34.95 (28.22)	24.96 (26.72)	9.99 (.00)*	381 (186/195)
Perceived Barriers to Sexual Health	2.85 (.55)	2.84 (.489)	.01 (.61)	349 (170/179)
Intention to Engage in Sexual Intercourse (% yes)	4.0%	2.0%	.02 (.20)	391 (194/197)
Intention to use Birth Control (% yes)	90.0%	87.0%	.03 (.12)	331 (160/171)

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

Note: Analysis controlled for the following variables: Age, Race, Pre-test scores, and the dummy coded imputed Pre-

Survey. \* = statistically significant at  $p \langle .01$ .

Table F.7. Post-intervention estimated effects using data from immediate post-survey (high attendance) to address the Secondary research questions

Outcome measure	Intervention mean or % (standard deviation)	Counterfactual mean or % (standard deviation)	Intervention compared to counterfactual mean difference (p-value of difference)	Sample Size (Intervention/ Counterfactual)
Sexual Activity Onset (% yes)	0.0%	0.0%	.00 (1.0)	522 (256/266)
Incidence of Pregnancy (% yes)	0.0%	0.0%	.00 (1.0)	522 (256/266)
Knowledge about STDs	33.47 (28.42)	16.19 (22.30)	17.28 (.00)*	484 (239/245)
Perceived Barriers to Sexual Health	2.83 (.58)	2.90 (.50)	06 (.17)	410 (204/206)
Intention to Engage in Sexual Intercourse (% yes)	1.0%	2.0%	01 (.99)	475 (235/240)
Intention to use Birth Control (% yes)	86.0%	83.0%	.03 (.26)	406 (197/209)

Source: Follow-up surveys administered immediately following the program.

Note: Analysis controlled for the following variables: Age, Race, Pre-test scores, and the dummy coded imputed Pre-

Survey. \* = statistically significant at  $p \langle .01$ .

Table F.8. Post-intervention estimated effects using data from six-month post-survey (high attendance) to address the Secondary research questions

Outcome measure	Intervention mean or % (standard deviation)	Counterfactual mean or % (standard deviation)	Intervention compared to counterfactual mean difference (p-value of difference)	Sample Size (Intervention/ Counterfactual)
Sexual Activity Onset (% yes)	0.0%	0.0%	.00 (1.0)	448 (218/230)
Incidence of Pregnancy (% yes)	0.0%	0.0%	.00 (1.0)	448 (218/230)
Knowledge about STDs	32.85 (29.05)	23.99 (27.71)	8.86 (.00)*	419 (205/214)
Perceived Barriers to Sexual Health	2.85 (.57)	2.86 (.51)	01 (.99)	372 (181/191)
Intention to Engage in Sexual Intercourse (% yes)	1.0%	0.0%	.01 (.08)	414 (205/209)
Intention to use Birth Control (% yes)	85.0%	87.0%	02 (.89)	356 (173/183)

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

Note: Analysis controlled for the following variables: Age, Race, Pre-test scores, and the dummy coded imputed Pre-

Survey. \* = statistically significant at  $p \langle .01$ .

## **APPENDIX G**

## IMPLEMENTATION (DOSAGE)

Table G.1. Dosage by Cohort

Sample	Attended 8 or more (Intervention/Counterfactual)	Attended Fewer than 8 (Intervention/Counterfactual)	
Cohort 01	57 (32/25)	24 (13/11)	
Cohort 02	148 (72/76)	64 (36/28)	
Cohort 03	225 (111/114)	69 (33/36)	
Cohort 04	150 (80/70)	59 (25/34)	
All	580 (295/285)	216 (107/109)	

# APPENDIX H PROGRAM SESSIONS

## Will Power/Won't Power

#### **Equal Earners, Savvy Spenders**

#### Introduction to Will Power/Won't Power

Program Overview Warm-up: ID Cards Intro to Relationships Romantic Relationships Conclusion/Reflection

#### Reproductive Health/Sexuality Review

Warm-up: What Am I? Female Health & Hygiene Myth Information Game Conclusion/Reflection

#### **Basic Assertiveness**

Five-Minute Check-in Intro to Assertiveness Practicing Assertiveness Conclusion/Reflection

#### **Identifying Sexual Pressures**

Five-Minute Check-in
Analyzing Media Messages
Ricky Rusiness

Risky Business Conclusion/Reflection

#### Looking at Values

Five-Minute Check-in Warm-up: Values Auction

Examining Values about Sexual Behavior

Have you Weighed your Options?

Conclusion/Reflection

#### The Case for Abstinence

Five-Minute Check-in Redefining Abstinence Don't' Let it Just Happen to You

Planning a Debate Conclusion/Reflection

#### **Resisting Sexual Pressure**

Five-Minute Check-in

Debating the subject of Abstinence Making Your Case: Pressure "Lines"

Practice Role Plays
Conclusion/Reflection

#### Defining your Decision: Look at the Risks

Five-Minute Check-in

The Pregnancy Probability Game

STD's: An Avoidable Risk

Evaluating the Risk of HIV Infection Letters to Lydia: Emotional Risks

Conclusion/Reflection

#### Sister Support System

Five-Minute Check-in
Exploring Sisterhood
Standing up for One Another
Pledge of Peer Support
Conclusion/Reflection

#### **Putting it all Together**

Five-Minute Check-in
Test Your Won't Power
Scripted Role Plays
Evaluation/Closure
Conclusion/Reflection
Parent-Daughter Workshop

#### Introducing Tomorrow's Equal Earners and Savvy Spenders

Money Association Mural Economy Puzzle Group Guidelines Conclusion Guidelines Conclusion/Reflection

#### **Getting What you Want and Need**

Welcome/Review

Wall of Wants and Needs

Working for Fun, Fulfillment, and Profit

Conclusion/Reflection

#### **Career Day**

Welcome/Review

Career Brainstorm: Three Sectors What are your Qualifications? Conclusion/Reflection

#### Bank on It!

Welcome/Review Banking Options

Help Balance my Checkbook Conclusion/Reflection

#### Being a "Loan Star"

Welcome/Review
The Price of Borrowing
Extra Credit
Conclusion/Reflection

#### Shop Smart

Welcome/Review

Building your Perfect Computer Pay Now or Pay Later? Conclusion/Reflection

#### Advertising and You

Welcome/Review Taking a Closer Look You Make the Ads Tell Them What you Think Conclusion/Reflection

#### **Know your Rights**

Welcome/Review
The Fairness Game
Conclusion/Reflection

# Taxes and Government Spending Welcome/Review

Tax Collector
Tax Freeze
Conclusion/Reflection

#### **Beyond Earning and Spending**

Welcome/Review

What do you Know? Matching Game

Celebrating the Whole You Conclusion/Reflection