

**Findings from an  
Innovative Teen  
Pregnancy  
Prevention  
Program**

# **Evaluation of the Be Yourself/Sé Tú Mismo in Maryland**

Final Impact Report for

George Washington University, Milken Institute School of Public Health

October 30, 2015

Prepared by

Amita N. Vyas, PhD  
Susan F. Wood, PhD  
Megan Landry, DrPH  
Grace Douglass, MPH  
Shelby Fallon, MPH

Recommended citation:

Vyas, A., Wood, S., Landry, M., Douglass, G., and Fallon, S. (2015). “The Evaluation of Be Yourself/Sé Tú Mismo in Montgomery & Prince Georges Counties, Maryland.” Washington, DC: The George Washington University Milken Institute School of Public Health.

Acknowledgements:

We would like to thank a number of people who made this evaluation possible. We recognize the staff of Identity, Inc. and Mary’s Center for their partnership in the successful implementation of the program and for serving as liaisons to the schools and youth. We thank the high schools and their administrators for collaborating and supporting this project. We acknowledge the many GW students that who helped with data collection, follow-up, and tracking. We also thank the many youth who participated in the evaluation and made this contribution to the field possible.

The Authors

This publication was prepared under Grant Number TP2AH000015 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 BE YOURSELF/ IN MONTGOMERY & PRINCE GEORGES COUNTY, MARYLAND: FINDINGS FROM AN INNOVATIVE TEEN PREGNANCY PREVENTION PROGRAM

## I. Introduction

The U.S. Department of Health and Human Services and the Office of Adolescent Health (OAH) is committed to improving the health and well-being of adolescents to enable them to become healthy, productive adults. A key component of OAH's mission is to support and evaluate evidence-based teen pregnancy prevention (TPP) programs. In 2010, the George Washington University Milken Institute School of Public Health, with its community partners Identity Inc. and Mary's Center, were funded through OAH's Tier 2 funding mechanism for implementing and evaluating new and innovative programs. The purpose of the project was to implement and evaluate the after school program, *Be Yourself/Sé Tú Mismo*, across twelve schools in Maryland to reduce risky sexual behaviors among a growing Latino youth population.

### A. Introduction & Study Overview

The Washington D.C. metropolitan region is a highly diverse urban center that has been a destination for an increasing number of immigrants and refugees from Central and South America,<sup>1</sup> reflecting a national demographic pattern that merits the attention of intervention research for these populations. Over the last decade, the number of Latinos in Montgomery County grew substantially. There are dense pockets of primarily immigrant communities within the region and Latinos can represent as high as 70 to 80 percent of the population.<sup>1</sup> Many of these families have limited English-speaking capabilities, which can pose an additional challenge for access to adequate social and health services, which are mostly provided in English.

The last decade has witnessed significant strides in the prevention of teen pregnancy. However, this overall trend masks the disproportionately high numbers of teen pregnancies experienced by minority groups in the United States, for whom prevention efforts have not been

as successful. Nationally, Latinos have higher teen birth rates than the overall U.S. population (26.5 births per 1,000 females ages 15-19 versus 41.7 teen births per 1,000 Latino females ages 15-19).<sup>2</sup> Although condom use at most recent sexual intercourse has been on the rise among adolescents (from 46 percent in 1991 to 59 percent in 2013), sexually active Latino adolescents were more likely than both white and black adolescents to have *not* used a condom or birth control during their last sexual intercourse.<sup>3</sup> In Montgomery County, Maryland, overall adolescent births are lower compared to national rates, yet Latino teen births are much higher (6.8 versus 22.9 per 1,000 births in 2012-2013).<sup>2</sup>

This report presents the methods and results of an impact and implementation evaluation of the *Be Yourself/Sé Tú Mismo* program in two counties in Maryland.

## **B. Primary Research Questions**

The impact evaluation's primary research questions are:

What is the impact of *Be Yourself/Sé Tú Mismo* compared to *Healthy Living/Vida Sana* on sexual debut six months after the program ends?

What is the impact of *Be Yourself/Sé Tú Mismo* compared to *Healthy Living/Vida Sana* on contraceptive use at last sex six months after the program ends?

What is the impact of *Be Yourself/Sé Tú Mismo* compared to *Healthy Living/Vida Sana* on contraceptive use in the last three months, six months after the program ends?

## **C. Secondary Research Questions**

The secondary research questions mirror the primary research questions and assess differences in all outcomes listed above immediately after the program ends.

## **II. Program & Comparison Programming**

Evaluation participants were offered either *Be Yourself/Sé Tú Mismo* or *Healthy Living/Vida Sana*. During the evaluation, both programs were implemented after-school in twelve high schools in Montgomery County or Prince George's Counties, Maryland. *Be Yourself/Sé Tú Mismo* is an after-school teen pregnancy prevention program using a cultural, developmental, and theory-based curriculum. *Healthy Living/Vida Sana* is an attention-control program that focuses on nonsexual health-related topics, such as fitness, nutrition, and exercise.

### **A. Description of Program as Intended**

*Be Yourself/Sé Tú Mismo* is an after-school program based on the Positive Youth Development model, a holistic strengths-based model which views youth as assets to be nourished rather than problems which need to be fixed. The Positive Youth Development model encourages the identification and strengthening of protective factors in youth's lives which will assist them in negotiating the risk factors they encounter on a daily basis. The program is designed to reduce teen pregnancy and other adverse risk behaviors among Latino youth. The program guides youth to think about and develop goals for their lives, and then think about possible obstacles, such as teen pregnancy, and how to avoid them. The program consists of four key components: 1) after-school sessions, 2) a weekend retreat, 3) social media outreach, and 4) Individual-Level Interventions/action plans and case management services. Refer to Appendix A for the program's logic model.

Youth Development Counselors (YDCs) from Identity, a community-based organization in Montgomery County that offers programs for Latino youth, delivered the program in 10 high schools in Montgomery County. YDCs from Mary's Center, a Federally Qualified Health Center that provides education, health and social services, delivered the program in two schools in Prince George's County. The program was offered in an after-school setting.

**The Curriculum:** The curriculum comprised a neutral module followed by five intervention modules. The neutral module, *Building Your Team*, focused on building a cohesive group, while allowing recruitment into the program to continue. This module included four sessions, which were identical for both the intervention and comparison programs. The remaining five intervention modules, which included 19 sessions, are entitled *You, Your Pit and Your Community*; *You and Your Emotions*; *You and Your Future*; *You and Your Relationships*; and *You and Your Goals*. Each intervention module addressed various aspects of the 5 Cs of Positive Youth Development: 1) competence—activities that are engaging and foster learning, 2) confidence—high expectations and clear limits for youth, 3) connection—meaningful opportunities for youth to engage in positive relationships with peers and adults, 4) caring and consistent relationships with adults, and 5) character—mechanisms to help youth move forward as young adults. Table II.1 lists the modules, the focus of each module, which of the 5 C’s of Positive Youth Development are covered, and the number of sessions.

**Table II.1. Description of the *Be Yourself/Sé Tú Mismo* Modules**

Module Name	Focus of Module	5 Cs of Positive Youth Development	Number of Sessions
<i>You, Your Pit and Your Community</i>	Discuss youth’s true selves and their personas	Competence, confidence, connection and caring	3
<i>You and Your Emotions</i>	Creates dialogue around being in tune with one’s own, and others’ emotions	Competence, confidence, character and caring	2
<i>You and Your Future</i>	Discuss locus of control, goal setting and future planning	Competence, confidence and connections	3
<i>You and Your Relationships</i>	Develop stronger connections to others, discuss different types of relationships and how to identify what youth they want from their relationships	Competence, confidence, connections, caring, and character	4

Module Name	Focus of Module	5 Cs of Positive Youth Development	Number of Sessions
<i>You and Your Retreat</i>	Reinforce objectives of the program outside of a regular after-school schedule, in a new, different, and relaxed environment	Competence, confidence, connections, caring, and character	2 + weekend retreat
<i>You and Your Goals</i>	Encourages youth to apply all that they have learned and discovered to themselves and their future goals	Competence, confidence, connections, caring, and character	5

YDCs implemented the 90-minute curriculum component sessions twice per week for 12 weeks. There were 23 sessions in cohort one and 19 sessions in the remaining five cohorts. Two sessions were eliminated due to the introduction of the neutral sessions and the need to allow more time for recruitment. Although two sessions were eliminated, the content was still delivered in the remaining sessions.

***The Weekend Retreat:*** YDCs led a weekend retreat, for each cohort, which allowed for 48 additional hours of curriculum programming and relationship building. In most cases, the retreat began Friday after school, included two overnights, and ended mid-day on Sunday. The retreats were held in retreat centers, in Western Maryland or on the Chesapeake Bay. This program component provided the opportunity to reinforce the objectives of the program outside of the regular after-school schedule, outside of the school building. Not only was this experience intended to provide a change to daily routine, but it also allowed the YDCs and youth to reach the program’s objectives in a different setting, focusing on teamwork and strengthening relationships between participants and program staff. The retreat participants also focused on creating individual long-term SMART (specific, measurable, achievable, relevant, and timely) goals, which then continued to be discussed in the remaining program sessions of the last module.

***Social Media Outreach:*** In addition to the in-person programming and support offered by program staff (described below), program participants received 12 pre-programmed text messages, delivered once per week through the Pro-texting platform. The messages were intended to reinforce information shared during the learning modules, and to remind participants of upcoming sessions or other program-related events. The messages were designed to align with the 5 Cs of Positive Youth Development. Each message addressed at least two of the 5 Cs, and was written using text communication in both English and Spanish that was familiar to the target population.

Study participants also joined specific Facebook groups that were created for each cohort to promote attendance, retention and a sense of community among members of the cohort. The same messages that were sent as text messages were also posted on Facebook each week. In addition, YDCs posted photos from sessions and the retreat on the Facebook pages, as well as other information. For example, if the session would be held outside the following day, students were reminded to bring warm clothes, or if the retreat was approaching, they were reminded to return their authorization forms.

***Individual Level Interventions and Case Management:*** Lastly, the program participants were provided Individual Level Interventions and built individual action plans with support from YDCs. Individual Level Interventions are one-on-one conversations between the YDC and a program participant, their family, or other interested party (i.e., school staff) whose purpose is to improve the participant's wellbeing. There is no required number of Individual Level Interventions, however YDCs generally met with youth for this purpose several times throughout the program. Participants required a varying number of Individual Level Interventions depending on their individual circumstances and progress towards their goals.



YDCs met with the family of each program participant. This meeting provided a time for the program staff to explain the program, answer any of the families' concerns about the program or the retreat, and to offer case management services. Both Identity and Mary's Center provide internal or external referrals for family planning, mental health and other healthcare referrals. Case management services also helped provide social services such as food and clothing assistance, emergency housing, legal aid, and employment opportunities for youth and their families.

## **B. Description of Counterfactual Condition**

The *Healthy Living/Vida Sana* Program is an attention-control program focused on nutrition, exercise and other non-sexual health topics. It has 3 core components: (1) group meeting sessions, (2) case management services, and (3) a weekend activity. Healthy Living Mentors, who were graduate students at George Washington University and Identity interns, led the program. There was no overlap between Healthy Living Mentors and YDCs.

***The Curriculum:*** Healthy Living Mentors implemented 12 90-minute sessions. The program was offered twice weekly for the first 8 sessions and once weekly for the remaining 8 sessions. The program began with the same four neutral sessions as the intervention group. The remaining sessions covered non-sexual health topics, such as nutrition and exercise.

***A Weekend Activity:*** The *Healthy Living/Vida Sana* program youth did not participate in an overnight weekend retreat, although they did have a daytime weekend activity as part of the program. The weekend activity was similar to a field day, and included team-building and physical activities, such as roller skating, bowling, and attending amusement parks. The activity lasted for about eight hours.

***Social Media Outreach:*** The comparison program also differed from the *Be Yourself/Sé Tú Mismo* intervention in that participants only received reminders to attend sessions through text messages and social media posts. No content was delivered via text or social media.

***Case Management:*** Participants in the *Healthy Living/Vida Sana* program also received case management services offered to the intervention group but did not receive Individual Level Interventions.

### **III. Study Design**

A cluster-randomized design, with matched pairs, was utilized for the evaluation. Cluster-randomized trials have two significant advantages including 1) increased efficiency for staffing/logistics, and 2) reduced risk of contamination (that is, of the comparison group receiving the intervention). A mixed-methods implementation study was also conducted, and included both quantitative and qualitative data. The following section describes the sample recruitment, study design, data collection, primary and secondary outcomes for analysis, study sample, baseline equivalence, and analytical methods.

#### **A. Sample Recruitment**

The population of interest was 9th- and 10th-grade self-identifying Latino students from 12 schools in Montgomery or Prince George's Counties, Maryland. Eligible high schools were selected based on the following: (1) a high percentage of Hispanic/Latino populations, and (2) a high percentage of students receiving free or reduced-price meals. Twelve eligible schools were recruited and all signed MOUs agreeing to participate over the 3 years of the intervention implementation period.

Youth were eligible to participate in the study if they:

- Attended 9th or 10th grade in the high school where the program was implemented
- Self-identified as Latino

- Spoke and/or understood Spanish since many program components were delivered in Spanish
- Were able to participate in one weekend activity
- Were able to attend two sessions per week on the days specified
- Were not pregnant at the time of recruitment (teen parents who were not pregnant were able to participate)
- Returned the completed application packet
- Had not participated in a previous *Be Yourself/Sé Tú Mismo* or *Healthy Living/Vida Sana* program through Identity or Mary's Center

Youth were recruited in schools through various methods such as, recruitment during lunch or in class, referrals from counselors or school staff, and in school announcements. Other recruitment methods included recruiting during out of school events such as PTA meetings or by sending home parent and student newsletters. During recruitment, interested youth were given application/enrollment packets. Youth volunteered to participate in the study, and youth were considered enrolled into the study if they completed an application packet (including parental consent) and attended at least one neutral session before randomization results were revealed.

## **B. Study Design & Random Assignment**

A cluster randomized design with matched pairs was used to estimate the impact of *Be Yourself/ Sé Tú Mismo*. The unit of random assignment was the school. Randomization for all cohorts was completed prior to implementation of cohort 1. Except for cohort 1 (prior to adding neutral sessions during the recruitment and enrollment period), schools were informed of their assignment *after* parental consent and collection of baseline survey data. For cohort 1, school assignment was shared with the project director and program manager prior to the completion of recruitment and enrollment. A greater number of youth enrolled in the intervention compared to the comparison for cohort 1, and it is plausible that is due to the early release of the random

assignment. As part of the sensitivity analysis described in Section III.G.1. impact analyses were conducted with cohorts 2-6.

Beginning with cohort 2, program sessions were modified to include four neutral sessions (identical for intervention and comparison schools) during the first two weeks of implementation. During this time, program staff continued to recruit youth for the program, which provided staff time to meet recruitment expectations while maintaining double blind recruitment. Consent forms did not indicate the schools' intervention status, and program staff were not aware of which school would receive which condition during recruitment. Randomization was shared with the program manager and program director on the last day of recruitment/neutral sessions. Youth were not allowed to enroll after random assignment results were shared.

Matched pairs were used to cluster schools. Schools were matched based on the following characteristics: (1) percentage of Hispanic/Latino students, (2) teen birth rate in school zip code, (3) miles from partner organization (up-county versus down-county location for Montgomery County schools only), and (4) school enrollment. One school in each matched pair was randomly assigned to receive the intervention and its matched school was randomly assigned to receive the comparison program. Randomization was restricted to ensure that six schools were assigned to the intervention and six schools were assigned to the comparison program for each cohort. The 12 schools were randomized for each of the six semester cohorts, yielding a final sample of 72 clusters.

### **C. Data Collection**

Impact evaluation data were collected via surveys for both intervention and comparison groups at four time points: baseline, immediate post-program, 6 months post-program, and 12

months post-program (the 12-month data was not analyzed for this report). Data on program implementation and fidelity were collected on an ongoing basis throughout the evaluation period.

## **1. Impact Evaluation**

Evaluation staff were responsible for all impact evaluation data collection for both intervention and comparison groups. Table B.1 in Appendix B presents an overview of the data collection schedule for the six cohorts. Baseline surveys were administered during the third and fourth neutral sessions prior to randomization results being shared and were collected in-person via mobile laptops with audio capabilities, and available in both English and Spanish. The immediate post-program follow-up survey occurred at the end of the program, and was collected during the last program session. The 6- and 12-month follow-up surveys were collected online via email or social media. If participants were unresponsive, telephone calls were made and surveys were administered by phone. Finally, surveys were mailed to participants who could not be reached by any of the aforementioned methods. There were no differences in the data collection protocol between the intervention and comparison schools. To maximize survey response rates and engagement in the study over time, participants received \$15 gift cards as incentives for completing surveys at each time point.

## **2. Implementation Evaluation**

The implementation evaluation entailed a mixed-methods approach with both qualitative and quantitative data collected from a variety of data sources. Data sources included 1) survey data on program satisfaction; 2) attendance logs; 3) implementation logs to assess fidelity for each session; 4) observations via the required OAH observation tool (Appendix: C) on 10% of randomly selected sessions each semester; and 5) in-depth interviews with staff. Youth in both intervention and comparison groups were asked a series of questions related to satisfaction

including: 1) overall program satisfaction and 2) satisfaction with the staff. These survey questions were measured on a Likert-type scale from 1-5 (strongly disagree to strongly agree), and reliability was also calculated for each scale used in the analysis. Additionally, youth were asked what they thought of the program overall (1=very poor to 5=excellent) and if the program should be taught to other youth their age (1=definitely, no; 5=definitely, yes).

Table D.1 in Appendix D provides an overview of the data sources used to evaluate implementation. All of the implementation data was collected for both the intervention and comparison groups.

#### **D. Outcomes for Impact Analyses**

Several outcomes on risky sexual behaviors were collected. Tables III.1 & III.2 describes the outcome variables for both primary and secondary research questions. Several variables required yes/no responses and were not re-coded. These dichotomous variables included 1) sexual debut; 2) contraceptive use in the last 3 months; and 3) contraceptive use at last sex.

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

Outcome name	Description of outcome	Timing of measure relative to program
Sexual Debut	<p>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:  “Have you ever had sexual intercourse?” This was defined as vaginal sex (when a males puts his penis into a female’s vagina)</p> <p><i>The variable is constructed as a dummy variable where respondents who respond yes, they have had sex are coded as 1 and those that have not had sex are coded as 0.</i></p>	6 months post-program

Outcome name	Description of outcome	Timing of measure relative to program
No contraceptive use in the last 3 months	<p>The variable is a yes/no measure of whether a person did not use contraceptives in the past 3 months if they had had sex. The measure is taken directly from the following item on the survey:</p> <p>“The next question is about your use of effective birth control methods. By effective methods, we mean the following:            Condoms (male and female)            Birth control pills            The shot (Depo Provera<sup>®</sup>)            The patch            The ring (NuvaRing)            IUD (Mirena or ParaGard)            Implant (Implanon)            In the <u>past 3 months</u>, have you had vaginal sex <u>without</u> you or your partner using any of these methods of birth control?”</p>	6 months post-program
<p><i>The variable is constructed as a dummy variable where respondents that have had sex in the past 3 months and did not use any of the listed contraceptive methods were coded as 1, and those that have had sex in the past 3 months but did use any of the methods were coded as 0. Respondents that did not have sex in the past 3 months were coded as skip-pattern missing (e.g., as 0).</i></p>		

Outcome name	Description of outcome	Timing of measure relative to program
Contraceptive use at last vaginal intercourse	<p>The variable is a yes/no measure of whether a person had used contraceptives the last time they had sex. The measure is taken directly from the following item on the survey: “Thinking about <u>the last time</u> you had vaginal sex, did you or your partner use any method of birth control?”</p> <p><i>The variable is constructed as a dummy variable where respondents that had ever had sex and used any method of birth control are coded as 1, and those that had ever had sex but did not use any form of birth control were coded as 0. Respondents that had never had sex were coded as skip-pattern missing.</i></p>	6 months post-program

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

Outcome name	Description of outcome	Timing of measure relative to program
Sexual debut	<p>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: “Have you ever had sexual intercourse?” This was defined as vaginal sex (when a males puts his penis into a female’s vagina)</p> <p><i>The variable is constructed as a dummy variable where respondents who respond yes, they have had sex are coded as 1 and those that have not had sex are coded as 0.</i></p>	Immediate post-program



Outcome name	Description of outcome	Timing of measure relative to program
No contraceptive use in the last 3 months	<p>The variable is a yes/no measure of whether a person did not use contraceptives in the past 3 months if they had had sex. The measure is taken directly from the following item on the survey:</p> <p>“The next question is about your use of effective birth control methods. By effective methods, we mean the following:            Condoms (male and female)            Birth control pills            The shot (Depo Provera <input type="checkbox"/>            The patch            The ring (NuvaRing)            IUD (Mirena or ParaGard)            Implant (Implanon)            In the <u>past 3 months</u>, have you had vaginal sex <u>without</u> you or your partner using any of these methods of birth control?”</p>	Immediate post-program
<p><i>The variable is constructed as a dummy variable where respondents that have had sex in the past 3 months and did not use any of the listed contraceptive methods are coded as 1, and those that have had sex in the past 3 months but did use any of the methods were coded as 0. Respondents that did not have sex in the past 3 months were coded as skip-pattern missing.</i></p>		

Outcome name	Description of outcome	Timing of measure relative to program
Contraceptive use at last vaginal intercourse	<p>The variable is a yes/no measure of whether a person had used contraceptives the last time they had sex. The measure is taken directly from the following item on the survey: “Thinking about <u>the last time</u> you had vaginal sex, did you or your partner use any method of birth control?”</p> <p><i>The variable is constructed as a dummy variable where respondents that have ever had sex and used any method of birth control are coded as 1, and those that have ever had sex but did not use any form of birth control were coded as 0. Respondents that have never had sex were coded as skip-pattern missing.</i></p>	Immediate post-program

## E. Study Sample

Table E.1 in Appendix E describes how the analytic sample was created for both the primary (6-month post-program) and secondary (immediate post-program) research questions. Due to the cluster design of the evaluation, the total sample size at the cluster-level is 72 (36 intervention clusters; 36 comparison clusters). At the individual youth-level, 1,356 youth consented to participate in the evaluation (707 intervention youth; 649 comparison youth), and 83% took the baseline survey (84% intervention; 82% comparison). For both the immediate post-program and 6-month post-program follow-ups, approximately 77% of the total sample completed the surveys. The response rate among the analytical sample between intervention and comparison at the 6-month follow-up was 79.5% for the intervention and 74.4% for the comparison. Thus, the differential attrition among the analytical sample between intervention and comparison at the 6-month follow-up was approximately 5 percentage points (21% intervention; 26% comparison). For the primary research questions, the final analytic sample is composed of youth who took

both the baseline and the 6-month follow-up survey (n = 911; 490 intervention; 421 comparison). The final analytic sample for the secondary research questions includes participants with data at both baseline and the immediate post-program follow-up (n = 912; 502 intervention; 410 comparison).

## F. Baseline Equivalence

Baseline equivalence tests for both the immediate post-program and the 6-month follow-up analytic samples were conducted to assess comparability of intervention and comparison groups at baseline. The baseline equivalency tables (Tables III.3a and III.3b) summarize baseline characteristics of youth by group, for the analytic samples, which consist of students who responded to the primary and secondary outcome measures. A multilevel, multivariate statistical model was utilized for the baseline equivalency. As shown, there are significant differences ( $p < .05$ ) between the intervention and comparison groups on key baseline characteristics. Therefore, the impact analyses adjusted for both baseline outcome variables and key demographic differences.

**Table III.3a. Summary statistics of key baseline measures for youth completing 6-month post-program survey**

Baseline measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison $p$ -value of difference adjusted for clustering
Age (in years)	15.34 (1.05)	15.51 (1.06)	-0.17	0.241
Grade (% 9th)	72.45	71.02	1.43	0.835
Gender (% male)	43.47	42.04	1.43	0.642
Survey language (% English)	71.63	66.03	5.6	0.321

Baseline measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison <i>p</i> -value of difference adjusted for clustering
U.S. born (% yes)	42.45	42.76	-0.31	0.699
Sexual Debut (% yes)	22.65	30.40	-7.75	0.035
No contraceptive use in last 3 months (% yes)	3.27	6.41	-3.14	0.048
Contraceptive use at last sex (% yes)	8.37	15.68	-7.31	0.012
Sample size	490	421	69	

**Table III.3b. Summary statistics of key baseline measures for youth completing immediate post-program survey**

Baseline measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison <i>p</i> -value of difference adjusted for clustering
Age	15.36 (1.11)	15.48 (1.08)	-0.12	0.398
Grade (% 9th)	72.71	70.73	1.98	0.770
Gender (% male)	45.62	43.90	1.72	0.668
Survey language (% English)	71.31	67.07	4.24	0.470
U.S. born (% yes)	42.03	43.41	-1.38	0.846
Sexual Debut (% yes)	23.90	29.27	-5.37	0.167
No contraceptive use in last 3 months (% yes)	3.39	5.12	-1.73	0.167

Baseline measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Intervention versus comparison mean difference	Intervention versus comparison <i>p</i> -value of difference adjusted for clustering
Contraceptive use at last sex (% yes)	9.96	14.63	-4.67	0.061
Sample size	502	410	92	

## G. Methods

To address the primary and secondary research questions, intention-to-treat (ITT) analyses were conducted. This approach estimates the impact of the program on all youth enrolled into the intervention or comparison group, regardless of the level of program participation. Furthermore, an ITT analysis is not subject to selection bias (i.e., the least committed youth will drop out of the intervention group or the comparison group), and the goal of the report is to meet HHS standards, and an ITT analysis is required for these standards.

### 1. Impact Evaluation

Multi-level linear probability modeling to estimate program impacts relative to the comparison program was conducted for both the primary and secondary analyses. The parameter estimate on the intervention variable in the regression models was interpreted as the impact of the intervention (see Appendix F for model specification). All models adjusted for clustering at the school level, and controlled for participant-level baseline characteristics (age, sex, survey language, and U.S. nativity), as well as baseline outcomes for each of the research questions. Results were adjusted for multiple comparisons using the Bonferroni correction to test each individual hypothesis at a significance level of  $\alpha/m$ . Thus, given there are three individual

hypotheses, the Bonferroni correction tests each individual hypothesis at a  $p$ -value of  $0.05/3 = 0.017$ .

Missing data arose for baseline, immediate, and 6-month post-program surveys. Pairwise deletion was used in the initial analyses. Further, dummy variable adjustment was also conducted to account for missing baseline covariates. Dummy variable adjustments did not reveal any differential results from the initial analysis, and are therefore, not included.

Sensitivity analyses were conducted to examine any statistical differences between various approaches. Sensitivity analyses consisted of analysis of 1) logistic regression models for the binary outcomes; 2) imputing positive responses for non-response missing data; 3) removing participant-level baseline covariates; and 4) conducting analyses for cohorts 2-6 only.

## **2. Implementation Evaluation**

The implementation study focused on several key domains for both the intervention and comparison program including: 1) adherence; 2) interest and engagement of youth; and 3) program satisfaction. The primary objective of the implementation analysis was to assess the degree to which the program was implemented as developed (fidelity) as well as participant recruitment, attendance, and satisfaction. Table G.1 in Appendix G provides details on the methods used to address implementation research questions.

## **IV. Study Findings**

The study findings focus on the impact of the intervention, relative to the comparison group on all primary outcomes, how the was program implemented and with what level of dosage for youth in both the intervention and comparison programs.

## **A. Implementation Study Findings**

Implementation analyses were conducted for both the 6-month post-program analytic sample and the immediate post-program analytic sample. The following describes results from the 6-month analytic sample (Table IV.1), however, results were similar for the immediate post-program analytic sample (Table IV.2).

For the 6-month post-program sample, all intended sessions were offered to both intervention and comparison groups. The program dosage for the curriculum component was 24.17 program sessions (there were 25 sessions for cohort 1 and then 19 sessions for cohorts 2-6), and 12 program sessions for the comparison group. Further, 93% (intervention) and 90% (comparison) of activities were completed. The program sessions were implemented as planned and with fidelity. However, as shown in Table IV.1, for the 6-month analytic sample, only 41.2% intervention youth and 32.7% comparison youth attended more than 75% of the sessions. Therefore, adherence (based on program session attendance) was much lower than expected with less than half of youth attending greater than 75% of sessions. For the retreat, 64.7% (n=325) of intervention youth (n=502) attended the retreat. Social media was recorded by how many Facebook pages and messages were created and sent, as well as how many text messages were sent for each cohort. Social media was also recorded by the percent of youth that joined the Facebook pages and the percent that signed up to receive text messages. Thirty-seven percent (n=336) of youth in the analytic sample signed up for Facebook, and 53% (n=482) joined the text message platform. In total, 72 Facebook pages were created for each of the 72 cohorts. There were, on average, 32 program content-specific messages that were sent via Facebook and text message per intervention cohort. Control cohorts received, on average, 10 logistic reminder messages to attend an upcoming session.

Youth were highly engaged and both youth and staff were highly satisfied with the programs, indicating high-quality implementation. For both the intervention and comparison programs, overall interest and engagement scores were high. However, youth in the intervention group were more interested (4.45 and 3.97,  $p < .001$ ) and engaged (4.43 and 3.96,  $p < .001$ ) in the program than youth in the comparison group. All program satisfaction scores were significantly higher for the intervention program compared to the comparison program. Youth in the intervention group thought the program should be taught to other students their age more than the youth in the comparison group (4.73 and 4.60,  $p < .001$ ). The staff scale (4.58 for intervention; 4.31 for comparison;  $p < .001$ ) and program satisfaction scale (4.58 for intervention; 4.29 for comparison;  $p < .001$ ) rated by youth yielded high scores.

**Table IV.1 Post-intervention estimated effects using data from 6-month post-program survey to address the implementation element**

Outcome measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Adjusted Treatment effect ( $p$ -value of difference)
Intended number of sessions	24.17	12	N/A
Proportion of sessions attended	64.75	61.79	0.1
Participants that did not attend any sessions (%)	2.82	0.99	N/A
Participants that attended >75% of sessions (%)	41.21	32.67	N/A
Proportion of activities completed (%)	93.00	90.00	0.024
Interest of participants	4.45 (.72)	3.97 (.95)	<0.001
Engagement of participants	4.43 (.73)	3.96 (.93)	<0.001
Overall, what did you think of this program?	4.55 (.75)	4.15 (.93)	<0.001
Should this program be taught to other students your age?	4.73 (.60)	4.60 (.72)	0.002



Outcome measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Adjusted Treatment effect ( <i>p</i> -value of difference)
Youth satisfaction with staff scale ( $\alpha=.954$ )	4.58 (.69)	4.31 (.78)	<0.001
Youth's program satisfaction scale ( $\alpha=.956$ )	4.58 (.69)	4.29 (.78)	<0.001
Sample Size	490	421	

Source: 6-month post-program survey data. *Follow-up surveys administered 5½ to 8 months after the program.*

Notes: The first cohort had 25 sessions for the intervention and all following cohorts had 19 sessions. The comparison group received 12 program sessions. Interest and engagement of participants was measured on the implementation logs filled out by staff. A 1-5 scale with 1 being the least interested/engaged and 5 being the most interested/engaged. Participants rated the overall program on a 1-5 scale (1=very poor to 5=excellent) follow-up surveys.

**Table IV.2 Post-intervention estimated effects using data from immediate post-program survey to address the implementation element**

Outcome measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Adjusted Treatment effect ( <i>p</i> -value of difference)
Number of sessions	24.17	12	N/A
Proportion of sessions attended	67.98	65.54	0.147
Participants that did not attend any sessions	1.69	0.25	N/A
Participants that attended >75% of sessions	45.45	37.15	N/A
Proportion of activities completed	93.00	90.00	0.024
Interest of participants	4.45 (.72)	3.97 (.95)	<0.001
Engagement of participants	4.43 (.73)	3.96 (.93)	<0.001
Overall, what did you think of this program?	4.45 (.80)	4.12 (.90)	<0.001
Should this program be taught to other students your age?	4.69 (.67)	4.36 (.93)	<0.001
Staff satisfaction scale ( $\alpha=.958$ )	4.57 (.69)	4.27 (.83)	<0.001

Outcome measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Adjusted Treatment effect ( <i>p</i> -value of difference)
Program satisfaction scale ( $\alpha=.954$ )	4.56 (.71)	4.27 (.80)	<0.001
Sample Size	502	410	

Source: Immediate post-program survey data. *Follow-up surveys administered immediately after the program.*

Notes: The first cohort had 25 sessions for the intervention and all following cohorts had 19 sessions. The comparison group received 12 program sessions. Interest and engagement of participants was measured on the implementation logs filled out by staff. A 1-5 scale with 1 being the least interested/engaged and 5 being the most interested/engaged. Participants rated the overall program on a 1-5 scale (1=very poor to 5=excellent) follow-up surveys.

## B. Impact Study Findings

There is no evidence that *Be Yourself/Sé Tú Mismo* caused statistically significant changes in any of the outcomes. Tables IV.3 and IV.4 present the impact analyses for the primary and secondary outcomes, respectively. *Be Yourself/Sé Tú Mismo* did not have an impact on whether youth ever had vaginal sex compared to *Healthy Living/Vida San* at either follow-up point. There were also no impacts on any of the contraceptive use measures at either follow-up point.

Sensitivity analyses consisted of analysis of 1) logistic regression models for the binary outcomes; 2) imputing positive responses for non-response missing data; 3) removing participant-level baseline covariates; and 4) conducting analyses for cohorts 2-6 only (Appendix H, Tables H.1 and H.2). There were no differences observed between the benchmark approach and the sensitivity approaches.

**Table IV.3. Post-intervention estimated effects using data from 6-month post-program survey to address the primary research questions**

Outcome measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Adjusted Treatment effect as percentage (p-value of difference)
Sexual Debut (% yes)	38.16	40.86	4.1 (0.109)
No contraceptive use in last 3 months (% yes)	6.12	5.94	-0.3 (0.838)
Contraceptive use at last sex (% yes)	16.94	18.29	3.0 (0.189)
Sample Size	490	421	

Source: 6-month post-program survey data. Follow-up surveys administered 5½ to 8 months after the program

Notes: All analyses control for age, gender, survey language, US nativity, baseline outcome measures, and adjust for clustering at the school level. See Table III.1 for a more detailed description of each measure and section III for a description of the impact estimation methods.

**Table IV.4. Post-intervention estimated effects using data from immediate post-program survey to address the secondary research questions**

Outcome measure	Intervention mean or % (standard deviation)	Comparison mean or % (standard deviation)	Adjusted Treatment effect as percentage (p-value of difference)
Sexual Debut (% yes)	31.27	32.44	0.2 (0.927)
No contraceptive use in last 3 months (% yes)	5.18	4.88	12.5 (0.090)
Contraceptive use at last sex (% yes)	13.75	15.37	4.8 (0.420)
Sample Size	502	410	

Source: Immediate post-program survey data. *Follow-up surveys administered immediately after the program.*

Notes: All analyses control for age, gender, survey language, US born, baseline outcome measures, and adjust for clustering at the school level. *See Table III.2 for a more detailed description of each measure and section III for a description of the impact estimation methods.*

## V. Conclusion

This study is one of the first rigorous evaluations of a Latino-focused youth development program aimed at reducing risky sexual behaviors. Although the findings indicated positive

changes on several sexual behaviors over time for both the intervention and comparison groups, *Be Yourself/ Sé Tú Mismo* did not have an impact on any of the sexual behaviors at either the immediate or 6-month post-program follow-ups.

It is important to note that although the curriculum component program was implemented with fidelity, participant attendance was low (64.75% for intervention youth and 61.79% for comparison youth), and it is plausible that this attributed to a lack of significant findings. An additional reason for the lack of significant findings may be that attending the after-school programs reduced the after-school hours where the participants could have engaged in risky sexual behavior. With adolescents, in particular, after-school hours are often times when they engage in risky sexual behavior. Therefore, regardless of the specific content, after-school programs that provide a structured environment with adult supervision are often effective at reducing risk-taking because the programming occurs in hours where youth could potentially engage in risky sexual behavior. Additionally, as the findings from the implementation analysis indicated, the *Healthy Living/Sana Vida* program was of high-quality and the comparison youth had positive changes in sexual risk-taking over time. Taken together, these two factors may have masked the differential impact of the intervention program.

Although the intention-to-treat analysis did not yield significant impact findings, this five-year research and demonstrated projected garnered significant lessons learned and provided the foundation for future Latino youth-based interventions.

First, youth participation in after school programs among low-income youth is particularly challenging given competing priorities (e.g., jobs, sports, caring for younger siblings). Therefore, after school programs need to take into the amount of time required for participation, and to devise alternate ways for participation (i.e. delivering content via through digital channels).

Second, the *Be Yourself/Sé Tú Mismo* program is based on decades of experience and research working with the Latino community. However, it is plausible that low attendance rates in the after school sessions compromised the results of the impact assessment. Therefore, the program should be adapted to increase attendance and then re-evaluated for impact.

Finally, the *Healthy Living/Sana Vida* program may have been successful at reducing risky behavior. It is plausible that the comparison program was well implemented and provided a safe and structured environment for youth, which reduced risky behaviors.

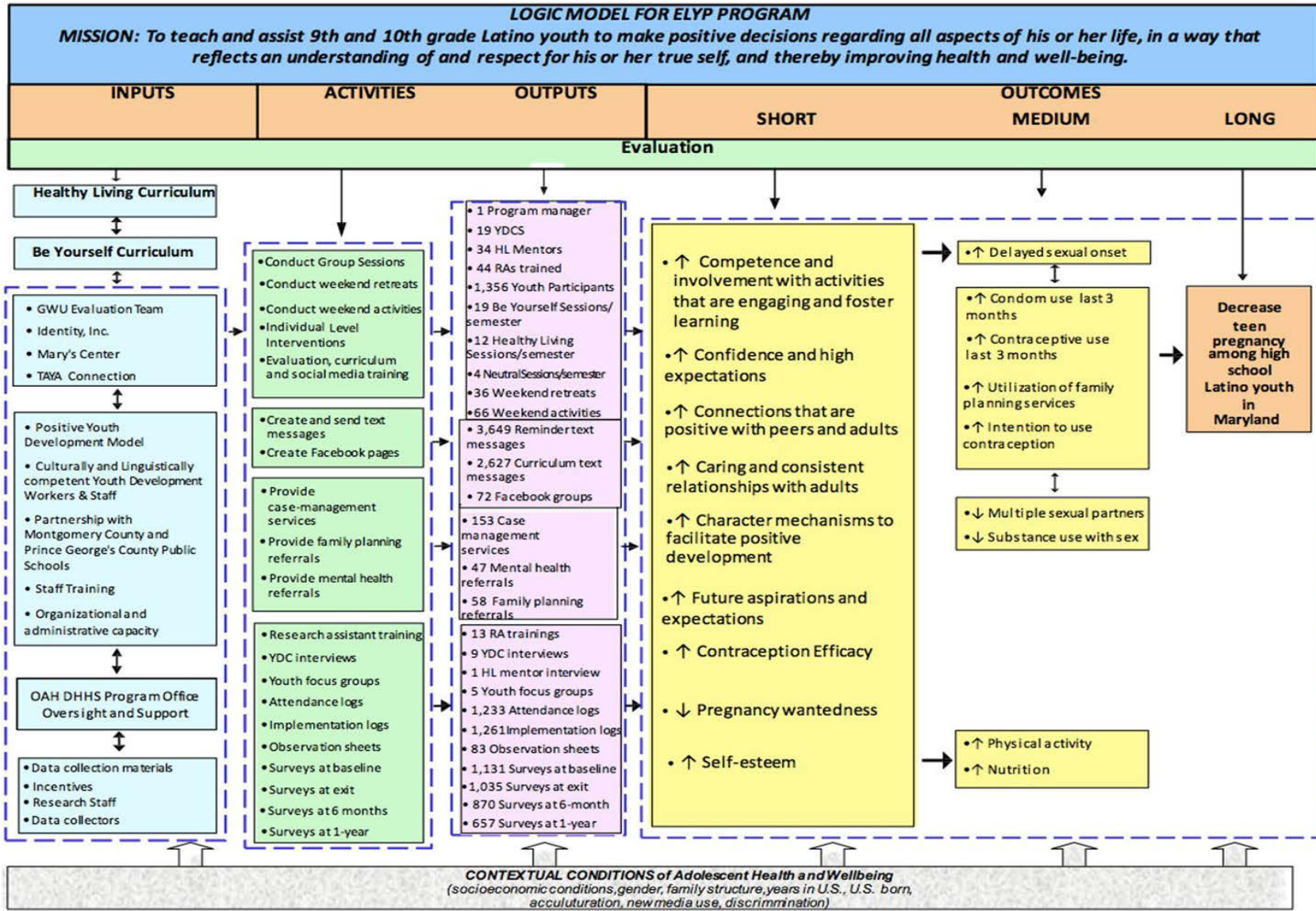
## References

<sup>1</sup> Pew Research Center Hispanic Trends (2011) *Demographic Profile of Hispanics in Maryland*, from <http://www.pewhispanic.org/states/state/md/>

<sup>2</sup> Office of Minority Health and Health Disparities Maryland Department of Health and Mental Hygiene. (2013) *Hispanics in Maryland: Health Data and resources*, from <http://dhmh.maryland.gov/mhhd/Documents/Maryland-Hispanic-Health-Disparity-Data.pdf>

<sup>3</sup> Centers for Disease Control and Prevention (CDC) (2014). 1991-2013 High School Youth Risk Behavior Survey Data. Accessed on 9/15/2014. Available at <http://nccd.cdc.gov/youthonline/>

# Appendix A: Logic Model



## Appendix B: Data collection efforts

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

Data collection effort	Cohort 1	Cohort 2	Cohort 3	Cohort 4	Cohort 5	Cohort 6
Start date of neutral sessions	9/19/2011	2/6/2012	9/19/2012	2/6/2013	9/23/2013	2/5/2014
Baseline survey	10/2011	2/2012	10/2012	2/2013	10/2013	2/2014
Program Implementation – 1 week after survey	10/2011	2/2012	10/2012	2/2013	10/2013	2/2014
Immediate post-program follow-up	12/2011	5/2012	12/2012	5/2013	12/2013	4/2014
6-month post-program follow-up	5/2012	11/2012	5/2013	11/2013	5/2014	11/2014



## Appendix C: OAH Observation Tool

### Program Observation Form for TPP Grantees

**Introduction:** The purpose of the observation form is to measure the fidelity and quality of implementation of the program delivery. Please use the guidelines below when completing the observation form and *do not* change the scoring provided; for example, do not circle multiple answers or score a 1.5 rather than a 1 or a 2.

**You should complete the observation form after viewing the entire session, but you should read through the questions prior to the observation.** It is also helpful to take notes during your viewing; for example, for Question 1, each time an implementer gives explanations, place a checkmark next to the appropriate rating.

**Instructions:** The following questions assess the overall quality of the program session and delivery of the information. Use your best judgment and do not circle more than one response.

#### 1. In general, how clear were the program implementer's explanations of activities?

1	2	3	4	5
Not clear		Somewhat clear		Very clear

- 1 - Most participants do not understand instructions and cannot proceed; many questions asked.
- 3 - About half of the group understands, while the other half ask questions for clarification.
- 5 - 90-100% of the participants begin and complete the activity/discussion with no hesitation and no questions.

#### 2. To what extent did the implementer keep track of time during the session and activities?

1	2	3	4	5
Not on time		Some loss of time		Well on time

- 1- Implementer does not have time to complete the material (particularly at the end of the session); regularly allows discussions to drag on (e.g., participants seem bored or begin discussing non-related issues in small groups).
- 3 - Misses a few points; sometimes allows discussions to drag on.
- 5 - Completes all content of the session; completes activities and discussions in a timely manner (using the suggested time limitations in the program manual, if available).

#### 3. To what extent did the presentation of materials seem rushed or hurried?

1	2	3	4	5
Not on time		Some loss of time		Well on time

- 1- Implementer doesn't allow time for discussion; doesn't have time for examples; tells participants they are in a hurry; body language suggests stress or hurry.
- 3 - Some deletion of discussion/activities; sometimes states but does not explain material.
- 5 - Does not rush participants or speech but still completes all the materials; appears relaxed.

**4. To what extent did the participants appear to understand the material?**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<b>Little understanding</b>		<b>Some understanding</b>		<b>Good understanding</b>

Use your best judgment based on participant conversations and feedback.  
Roughly: 1 - Less than 25% seem to understand; 3 - About half; 5 - 75-100% understand.

**5. How actively did the group members participate in discussions and activities?**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<b>Little participation</b>		<b>Some participation</b>		<b>Active participation</b>

Use your best judgment based on listening to the discussions and feedback.  
Roughly, 1 - Less than 25% participate; 3 - About half participate. 5 - 75-100% participate

**6. On the following scale, rate the implementer on the following qualities:**

**a) Knowledge of the program**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<b>Poor</b>		<b>Average</b>		<b>Excellent</b>

1 - Cannot answer questions, mispronounces names; reads from the manual.  
5 - Provides information above and beyond what's in the manual; seems very familiar with the concepts and answers questions with ease.

**b) Level of enthusiasm**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<b>Poor</b>		<b>Average</b>		<b>Excellent</b>

1 - Presents information in a dry and boring way; lacks personal connection to material; appears "burned out."  
5 - Makes clear that the program is a great opportunity; gets participants talking and excited; outgoing.

**c) Poise and confidence**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<b>Poor</b>		<b>Average</b>		<b>Excellent</b>

1 - Appears nervous or hurried; does not have good eye contact.  
5 - Does not hesitate in addressing concerns. Well organized, not nervous.

**d) Rapport and communication with participants**

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<b>Poor</b>		<b>Average</b>		<b>Excellent</b>

1 - Doesn't remember names; does not "connect" with participants; acts distant or unfriendly.  
5 - Gets participants talking and excited; very friendly; uses people's names when appropriate; seems to understand the community and its needs.

**e) Effectively addressed questions/concerns**

**1**  
**Poor**

**2**

**3**  
**Average**

**4**

**5**  
**Excellent**

- 1 - Engages in "power struggles"; responds negatively to comments; gives inaccurate information; doesn't direct participants elsewhere for further info.
- 5 - Answers questions of fact with information, questions of value with validation; if doesn't know the answer, is honest about it and directs them elsewhere.

**7. Rate the overall quality of the program session.**

**1**  
**Poor**

**2**

**3**  
**Average**

**4**

**5**  
**Excellent**

Summary measure of all the preceding questions. Assesses both the extent of material covered and the performance of the implementer.

Excellent sessions looks like:

- Participants are doing rather than talking about activities
- Non-judgmental responses to questions
- Answering questions of fact with information, questions of value with validation
- Good time management and well organized
- Adequate pacing—not too fast and did not drag
- Using effective checks for understanding.

Poor sessions look like:

- Lecture-style of presenting the content
- Reading the content from the notebook
- Stumbling along with the content and failing to make connections to what has been discussed previously or what participants are contributing.
- Uninvolved participants
- Getting into power struggles with participants about the content.
- Judgmental responses
- Flat affect and boring style
- Unorganized and random
- Loses track of time.

Note: The following questions (8, 9, and 10) are for grantee's internal use only for program improvement purposes. These questions are optional and will not be reported to OAH or ACYF for performance measurement purposes.

**8. Briefly describe any implementation problems you noticed, including any major changes to the content or delivery of the material; time wasted in getting the session started or finished, etc:**

---

---

---

**9. Please note at least one major strength of the session and/or facilitator's delivery of the material:**

---

---

---

**10. Other Comments: Use the space below for additional comments regarding strengths or**

**weaknesses of the session, particularly if there is anything that affected your ratings above.**

---

---

---

## Appendix D: Implementation evaluation data collection

**Table D.1. Data used to address implementation research questions**

Implementation Element	Data Source(s)	Frequency/sampling of data collection	Person(s) Responsible
<b>Adherence</b>			
How many and how often were sessions offered (for example, number of sessions delivered, average duration, average frequency)?	Intervention protocol Implementation logs	After each program session	YDCs, Healthy Living Mentors Program Manager
What and how much was received (for example, average number and percent of sessions attended, percentage of sample that did not attend at all [no-shows])?	Attendance logs (paper and electronic)	After each program session	YDCs, Healthy Living Mentors
What content was delivered to youth (for example, total number of topics covered, proportion of material that was ultimately discussed in sessions, proportion of content delivered by session & module)?	Implementation logs	After each program session	YDCs, Healthy Living Mentors
	Observation Tool	Observed 10% of sessions	Evaluation team

Implementation Element	Data Source(s)	Frequency/sampling of data collection	Person(s) Responsible
<b>Quality</b>			
Quality of youth engagement with program	Observation Tool	Observed 10% of sessions	Evaluation team
	Implementation Log	After each program session	YDCs, Healthy Living Mentors
	Staff interviews	Twice a year	Evaluation team
	Program satisfaction questions on surveys	Immediate post program and 6-month post-program surveys	Evaluation team

## Appendix E: Study Sample

**Table E.1. Cluster and youth sample sizes by intervention status**

Number of:	Time period	Total sample size	Intervention sample size	Comparison sample size	Total response rate %	Intervention response rate %	Comparison response rate %
Clusters: At beginning of study		72	36	36			
Clusters: Contributed at least one youth at baseline	Baseline	72	36	36	100	100	100
Clusters: Contributed at least one youth at follow-up	Immediate post-program	72	36	36	100	100	100
Clusters: Contributed at least one youth at follow-up	6 months post-program	72	36	36	100	100	100
Youth: In non-attriting clusters/sites at time of assignment		1,356	707	649			
Youth: Who consented		1,356	707	649	100	100	100
Youth: Contributed a baseline survey	Baseline	1,131	596	535	83.4	84.3	82.4
Youth: Contributed a follow-up survey	Immediate post-program	1,038	565	473	76.5	79.9	72.9
Youth: Contributed a follow-up survey	6 months post-program	1,046	562	483	77.1	79.5	74.4

## Appendix F: Model Specification

Individual outcomes were modeled at level 1, and level 2 represented the level cluster randomization at the school level.

$$\text{Level 1: } Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \varepsilon_{ij}$$

$$\text{Level 2: } \beta_{0j} = \gamma_0 + \gamma_1T_j + \gamma_sD_{sj} + \mu_j$$

Where at Level 1:

$Y_{ij}$  is the outcome on the dependent variable for individual  $i$  in cluster  $j$ .

$\beta_{0j}$  is the intercept of the outcome variable in cluster  $j$  (Level 2).

$\beta_{1j}$  is the slope for the relationship in group  $j$  (Level 2) between the Level 1 predictor and the outcome for individual  $i$  in cluster  $j$ .

$X_{ij}$  refers to the Level 1 predictor for individual  $i$  in cluster  $j$

$\varepsilon_{ij}$  is the random error for the Level 1 equation for individual  $i$  in cluster  $j$ , and is assumed to be independently and identically distributed.

Where at Level 2:

$\gamma_0$  is the overall intercept. This is the grand mean of the scores on the dependent variable across all the groups when all the predictors are equal to 0.

$\gamma_1$  is the coefficient of interest, which represents the estimated impact of the intervention.

$T_j$  is a dummy variable equal to 1 if assigned to the intervention group.

$D_{sj}$  are dummy variables representing the randomization strata, school.

$\mu_j$  is the random error for school  $j$ , which is assumed to have a mean of zero, and independent from the errors at the individual level.



## Appendix G: Implementation evaluation methods

**Table G.1. Methods used to address implementation research questions**

Implementation Element	Methods Used to Operationalize Each Element
	<b>Adherence</b>
How many and how often were sessions offered (for example, number of sessions delivered, average duration, average frequency)?	<p>The total number of sessions is a sum of the sessions captured in implementation logs</p> <p>Average session duration is calculated as the average of the observed sessions lengths, measured in minutes</p> <p>Average weekly frequency is calculated as the total number of sessions divided by the total number of weeks when programming was offered</p>
What and how much was received (for example, average number and percent of sessions attended, percentage of sample that did not attend at all [no-shows])?	<p>Average number of sessions that each participant attended</p> <p>Percentage of sessions is calculated as the total number of sessions attended divided by the total number of sessions offered</p> <p>Percentage of youth who attended at least 75% or more of program sessions</p>
What content was delivered to youth (for example, total number of topics covered, proportion of material that was ultimately discussed in sessions)?	<p>Average number of activities completed by program session</p> <p>Total number of activities completed divided by the sum of total number of activities (calculated for each program session; calculated by total activities across all program sessions)</p>

Implementation Element	Methods Used to Operationalize Each Element
<b>Quality</b>	
Quality of youth engagement with program	<p>Percent of observed sessions that scored high or very high for youth engagement</p> <p>Mean score on youth participation and understanding (2 questions) and overall program quality (1 question) from program observation tool</p> <p>Mean score on youth engagement from implementation logs</p> <p>Descriptive qualitative themes from YDC interviews questions about engagement</p>
<b>Satisfaction</b>	
Youth satisfaction with program	<p>Mean score on youth satisfaction with staff</p> <p>I enjoyed being in this program</p> <p>I had fun in this program</p> <p>I feel the program was useful for me</p> <p>I felt safe in the program</p> <p>I would recommend the program to my friends</p>
Youth satisfaction with staff	<p>Mean score on youth satisfaction with staff</p> <p>I felt the staff listened to me</p> <p>I felt the staff were open and honest</p> <p>I felt the staff were concerned about my issues</p> <p>I felt the staff included all participants in the sessions</p> <p>If I had a problem or question, the staff were always available to talk with me</p> <p>I felt the staff were respectful towards everyone during the program</p>

## Appendix H: Sensitivity analyses

The first sensitivity analysis tested whether a logistic regression model produced comparable results to the linear probability model, due to impacts on dichotomous outcomes estimated with a linear probability model for ease of interpretation. The second analysis tested whether the benchmark findings were replicated when missing responses between baseline and follow-up surveys were set to positive (e.g., missing follow-up data on “Ever had Sexual Intercourse” was imputed to a ‘yes’ response). The third analysis assessed the effect of removing individual-level baseline covariates in the impact model. Due to concerns about the random assignment regarding Cohort 1, and if it was truly random, the final analysis assessed the effect of excluding Cohort 1 in the analysis.

**Table .1. Sensitivity of impact analyses using data from 6-month post-program survey to address the primary research questions**

Condition 1 compared to Condition 2	Benchmark approach impact (SE)	Benchmark approach p-value	Positive responses for missing (SE)	Positive responses for missing p-value	Logistic (SE)	Logistic p-value	Removing covariates (SE)	Removing covariates p-value	Cohorts 2-6 Only (SE)	Cohorts 2-6 Only p-value
Sexual Debut	4.1 (.023)	0.109	.41 (.023)	0.109	50.7 (.370)	0.095	-2.7 (.028)	0.354	05.6 (.024)	0.041
No contraceptive use in last 3 months	-0.3 (.016)	0.838	-2.1 (.033)	0.546	-10 (.307)	0.759	-0.1 (.018)	0.943	0.5 (.015)	0.737
Contraceptive use at last sex	3.0 (.022)	0.189	-0.05 (.031)	0.988	28.3 (.234)	0.171	-2.0 (.023)	0.414	4.1 (.021)	0.080

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

**Table H.2. Sensitivity of impact analyses using data from immediate post-program survey to address the secondary research questions**

Condition 1 compared to Condition 2	Benchmark approach difference (SE)	Benchmark approach difference p-value	Positive responses for missing (SE)	Positive responses for missing p-value	Logistic (SE)	Logistic p-value	Removing Covariates (SE)	Removing covariates p-value	Cohorts 2-6 Only (SE)	Cohorts 2-6 Only p-value
Sexual Debut	0.2 (.016)	0.927	-3.0 (.025)	0.262	-1.6 (.233)	0.980	-4.4 (.033)	0.,206	0.9 (.017)	0.633
No contraceptive use in last 3 months	12.5 (.067)	0.090	2.6 (.034)	0.454	97.5 (.733)	0.067	12.8 (.070)	0.095	15.4 (.075)	0.064
Contraceptive use at last sex	4.8 (.057)	0.420	-0.03 (.041)	0.994	25.3 (.342)	0.408	-0.04 (.070)	0.995	6.4 (.072)	0.396

Source: *Follow-up surveys administered six months post-program*