Final Impact Evaluation Report

Early Findings from the Evaluation of the Pono Choices Program—A Culturally Responsive Teen Pregnancy and Sexually Transmitted Infection Prevention Program for Middle School Youth in Hawai‘i

March 11, 2016

Submitted to:
University of Hawai‘i at Mānoa,
Center on Disability Studies

Submitted by:
IMPAQ International, LLC
10420 Little Patuxent Parkway, Suite 300
Columbia, MD 21044

Authors:
Yasuyo Abe (Impact Study Lead), Linda Toms Barker (Evaluation Director), Vincent Chan (Senior Research Analyst), and Jasmine Eucogco (Senior Research Analyst)
IMPAQ International

Neither IMPAQ International LLC nor any contributors to this report have financial interests that could be affected by findings from the evaluation of Pono Choices. Throughout the study, the IMPAQ study team, as a third-party evaluator, worked independently of and separately from the University of Hawai‘i staff.

This publication was made possible by Grant Number TP2AH000017 from the Office of Adolescent Health, U.S. Department of Health & Human Services.

ACKNOWLEDGEMENTS

The authors thank the principals, teachers, and students who participated in the data collection for their cooperation and contribution to this study.

The authors thank the University of Hawai'i team—especially Kelly Roberts, Tammy Tom, Holly Manaseri and Denise Uehara—for their support for the evaluation and their sharing key implementation data. We thank the entire Implementation Team (Pono Choices Program Staff) led by Kelly Roberts, Ph.D., Principal Investigator, and staff of the Center on Disability Studies, Planned Parenthood of Hawai'i and ALU LIKE, Inc., for their tireless efforts to implement the program within the parameters required to support the rigorous evaluation design.

The authors also thank current and former staff of IMPAQ International who contributed to the collection of data and preliminary data analyses, including Futoshi Yumoto, Colleen McLelland, Paola Rochabrun, Kristen Bard, and Hui Kim, as well as our excellent subcontractor, Ward Research, who contributed to the collection of data, including Angela Llana, Leanne Nakamura, and Margarita Ayala.

TABLE OF CONTENTS

LIST OF EXHIBITS

I. INTRODUCTION

I.A. Introduction and Study Overview

This report presents findings from an impact study of the Pono Choices teen pregnancy and sexually transmitted infection (STI) prevention curriculum, funded with a Teen Pregnancy Prevention (TPP) Research and Demonstration Program Grant from the Office of Adolescent Health, U.S. Department of Health and Human Services. The Pono Choices program was developed by the University of Hawai‘i at Mānoa (UH Mānoa), in partnership with Planned Parenthood of Hawai‘i and ALU LIKE, Inc. The impact study was conducted by IMPAQ International, LLC, an evaluator contracted by UH Mānoa. The IMPAQ evaluation team conducted data collection and analysis independently of the program implementation team led by UH Mānoa.

Pono Choices was developed in 2011, as the State of Hawai‘i faced a continuing need for effective reproductive health education for its youth. Hawai‘i had the lowest rate of condom use among U.S. high school students in 2011 (44 percent, compared to 60 percent nationwide) and the tenth highest teen pregnancy rate in 2010 (6.5 percent, compared to 5.7 percent nationwide) (Eaton, Kann, Kinchen, et al., 2011; Kost & Henshaw, 2014). Furthermore, minority females in this diverse state (25 percent white, 39 percent Asian, 10 percent Native Hawaiian/Pacific Islander, and 24 percent mixed races) were disproportionately at high risk of teen births (29 births per 1,000 Asian/Pacific Islanders aged 15-19, compared to 22 births per 1,000 white counterparts in 2012) (U.S. Census Bureau, 2010; Hamilton, et al., 2014). The state’s chlamydia rate had been consistently higher than the national average over the previous decade (2001-10), and had only recently improved (8.5 percent, compared to 8.4 nationwide in 2011) (Centers for Disease Control and Prevention, 2015).

The goal of Pono Choices is to equip middle school youth with the knowledge and skills necessary to reduce their risk of unintended pregnancy and STIs by providing medically accurate information with instructional strategies that emphasize the Hawaiian host culture (see Appendix A for additional information on Hawaiian host culture). When developed in 2011, the curriculum met the State of Hawai‘i Abstinence-Based Policy (State of Hawai‘i Board of Education, 1995) and its middle school sexual health and responsibility standards and benchmarks (Hawaii Department of Education, 2007).

The Pono Choices curriculum, targeting youth ages 11 thru 13, is based on a prevention framework that emphasizes intervention in early adolescence. Moore and Sugland (2001) describe research showing that negative behavior patterns among adolescents have their origins in childhood, and recommend that interventions start before puberty, particularly for at-risk youth from disadvantaged or dysfunctional families. The Pono Choices' approach is also supported by research that suggests the best way to prevent high-risk behaviors that often precede sexual activity, such as drug and alcohol use, is to work with youth in early adolescence (Moore & Sugland, 2001; Hawkins, Catalano, Kosterman, Abbott & Hill 1999).

Pono Choices is unique in that it is developed exclusively for youth in Hawai'i, emphasizing Hawaiian cultural values and practices to promote positive character development. Curriculum that is responsive to culture "is an approach that empowers students intellectually, socially, emotionally, and politically by using cultural referents to impart knowledge, skills, and attitudes" (Ladson-Billings, 1995). Currently, no other current state-approved curriculum explicitly incorporates the unique values and perspectives of the Hawaiian host culture (Hawai'i Department of Education, 2015). (See Appendix A for additional information on Hawaiian cultural values and practices and on place-based curricula focusing on host culture.)

To assess the impact of Pono Choices, a cluster randomized controlled trial was conducted, in which the school was the unit of assignment. The study recruited middle schools across the state of Hawai‘i and included in the sample 7th- and 8th-grade students who were enrolled in the target health education classes. The study was originally designed to assess the impact of Pono Choices on behavioral outcomes two years after the intervention, when students reach the age when they are more likely to be involved in sexual relationships and have occasions to apply the knowledge and skills gained through the intervention. Due to unexpected changes in Hawai'i Department of Education's research restrictions, it was not possible to complete the 2-year follow-up survey data collection. The current report shares findings based on the 1-year follow-up survey, examining the effects of Pono Choices on sexual behaviors approximately 1 year after baseline.

I.B. Primary Confirmatory Research Question

The primary hypothesis investigated in this study was that students who participate in the Pono Choices curriculum would be less likely to engage in high-risk sexual behaviors than students in the control group 1 year after baseline: *What was the impact of Pono Choices, relative to business as usual, on engagement in high- risk sexual behavior (defined as intercourse with neither a condom nor effective method of birth control) among youth as measured 1 year after baseline?*

I.C. Secondary Confirmatory Research Question

The study examined another hypothesis that students who participate in the Pono Choices curriculum will be more likely to delay initiation of sexual activity than students in the control group at 1-year after baseline: *What was the impact of Pono Choices, relative to business as usual, on the initiation of sexual activity among youth as measured 1 year after baseline?*

I.D. Exploratory Research Questions

In addition to the primary behavioral outcomes, the study explored potential effects on the non-behavioral outcomes Pono Choices was designed to affect: knowledge, attitudes, skills, and intentions. Given the young age of program participants (average age of 12 at baseline), the majority of study participants were expected to remain abstinent over the follow-up period. For example, Centers for Disease Control and Prevention (2015) reports that only 10 percent of youth in Hawai‘i had ever had sex by 8th grade in 2013. With the majority expected to be not yet sexually active, distal behavioral outcomes may not be fully affected just one year after baseline. To gain insights into potential early effects of the intervention, the study examined the proximal non-behavioral outcomes at the 1-year follow-up.

According to the logic model of Pono Choices (Appendix B), these non-behavioral outcomes are expected to have a bearing on healthy sexual behavior later in adolescence. The study does not formally test the mediating effects of non-behavioral factors on behavioral outcomes, and evidence supporting a causal relationship between behavioral and non-behavioral measures in the field is scant. However, previous correlational studies suggest that these non-behavioral outcomes may be linked to changes in behavior. For example, studies have found that knowledge of contraceptive methods was correlated with higher odds of contraception use, and lower odds of not using any method of contraception (Frost, Lindberg, & Finer, 2012). Another study found that young adults who believed that avoiding pregnancy was not very important were more likely to become pregnant in the following year (Kornides, Kitsantas, Lindley, & Wu, 2015). The exploratory questions in this study are: *What was the impact of Pono Choices, relative to business as usual,*

. . . on students' knowledge of pregnancy and STI prevention at 1 year after baseline?

. . . on students' attitudes toward healthy sexual behavior at 1 year after baseline?

. . . on students' skills in managing relationships and choices at 1 year after baseline?

. . . on students' intentions about future sexual activity at 1 year after baseline?

II. PROGRAM AND COMPARISON PROGRAMMING

II.A. Description of Program as Intended

As described in Section I, Pono Choices was designed for middle school youth and emphasizes medically accurate content and the values and practices of the Hawaiian host culture. The logic model for the intervention (Appendix B) illustrates that youth exposed to Pono Choices are expected to improve their knowledge, attitudes, skills, and intentions toward pregnancy and STI prevention, which will then lead to delayed initiation of sex and a reduction in high-risk behavior in later years. Pono Choices is intended to be implemented in settings where youth ages 11-13 congregate for sexual health education, such as schools, community centers, or health clinics. This study tested the effectiveness of the Pono Choices curriculum in a middle school setting as part of regular health instruction for 7th- and 8th-graders.

The Pono Choices curriculum was designed to address the Hawai'i standards for sexual health education in middle schools regarding teen pregnancy and STI prevention. Pono Choices was developed as an alternative to currently available curricula, rather than as a supplement. The core contents of the Pono Choices curriculum consist of a set of specific knowledge components, attitudes, and skills derived from a review of the evidence-based teen pregnancy and STI prevention programs. Knowledge development focuses on the causes, transmission, and prevention of STIs; reproductive anatomy; and birth control. Attitude development addresses beliefs and opinions about abstinence, STIs, and pregnancy. Skills development focuses on negotiation and refusal skills, the correct procedures for condom use, and increased self-efficacy.

The contents is delivered in three stages. The first stage focuses on setting a foundation of trust and basic knowledge—establishing group agreements, identifying goals, identifying support networks, examining healthy relationships, understanding normal developmental changes during puberty, and knowing definitions of sex and abstinence used in the curriculum. The second stage focuses on knowledge about birth control and the causes, transmission, and prevention of STIs. It also covers beliefs and attitudes about abstinence, STIs, and teen pregnancy. The third stage focuses on building effective decision-making and negotiation skills and on increasing self-efficacy as students demonstrate knowledge of prevention skills through role play and practice.

The curriculum uses an instructional manual and a program kit that includes slides, videos, audio recordings, and other materials needed for each lesson. The curriculum includes 9.5 hours of content divided into 10 sequential modules. The first, introductory module is 30 minutes, and the remaining nine modules are 60 minutes each (see Appendix A for a summary of each module). The basic delivery model assumed 10 sessions (one module delivered per session), but the content can be delivered in fewer or more sessions, depending on the length of the sessions.

The Pono Choices curriculum was designed to accommodate the potential range of teacher skill levels. Teachers are trained to use the scripted materials through a 2-day training prior to implementation. In addition, teachers receive on-going support, including observation and feedback from instructional coaches and a 1-day (6-hour) refresher training each semester.

II.B. Description of the Counterfactual Condition

The counterfactual condition in the participating control group schools was the business-as-usual sexual health instruction that takes place in regular 7th and 8th grade health classes. The State of Hawai'i requires a sexual health education program that is age-appropriate and medically-accurate but does not use a standard curriculum across schools. The Hawai‘i Department of

Education does not prescribe specific levels of dosage or contents for schools, but provides a list of authorized curricula. While the department has identified specific approved curricula that schools can choose from, schools have been free to choose how much of the material they use, without any requirement that they implement a complete curriculum. Schools are also free to bring in outside experts and presenters to teach the material, teach it themselves, or both.

III. STUDY DESIGN

III.A. Sample Recruitment[1]

All regular public and public charter schools were eligible to participate in the study except for regular public middle schools that included 6th grade (because they were being considered for a different study). Participation in the study required that the school commit to random assignment to either the intervention or control group. Prior to randomization, schools were also required to identify one health or physical education teacher who taught 7th and/or 8th grade health classes and agreed to cooperate in the study. If the school was assigned to the intervention group, the teacher who agreed to cooperate with the study would implement Pono Choices as the sexual health curriculum. If assigned to the control group, the teacher who agreed to cooperate with the study would continue with the school's business-as-usual health curriculum.

Schools were recruited in two rounds—in Fall 2011 and in Spring 2012—from O'ahu, Big Island, Maui, Moloka'i, and Kaua'i. Recruitment efforts focused on the 83 middle schools (54 regular public schools and 29 public charter schools) in the state of Hawai'i, which operates a single statewide public school district. The implementation team led the recruitment efforts, with the evaluation team assisting in explaining evaluation activities. Overall, 36 schools agreed to

[1] See Appendix D for additional information on recruitment efforts.

participate in the study (19 in Round 1 and 17 in Round 2), including 21 regular public schools, 14 public charter schools, and one private school.[2]

III.B. Design

The study was designed as a cluster randomized controlled trial, with the school as the unit of randomization. The 36 middle schools that agreed to participate in the study were randomized after they committed to the terms of the study, confirmed the participation of one health education teacher, and signed a memorandum of understanding about study participation. Round 1 schools recruited during Fall 2011 were randomly assigned prior to the start of Spring 2012 when the implementation period began, and Round 2 schools recruited during Spring 2012 and randomized prior to Fall 2013. Schools were blocked by island (O‘ahu vs. neighbor islands) and type of school (regular public vs. public charter/private). In addition, timing of planned implementation of the health curriculum (early vs. later in the semester) was used to block schools in Round 1. Timing was not used in Round 2, as we discovered that the teacher-reported planned timing did not accurately reflect when teachers actually started the sexual health unit. Also, given the variation in the number of weeks needed to complete the curriculum (depending on length of sessions and number of sessions per week) in the second and third semesters, teachers were encouraged not to wait until late in the semester to implement.

Random assignment yielded 18 intervention and 18 control schools. Immediately after randomization, one public charter intervention school and one regular public control school left the study and declined further contact, and no additional information was collected from them. The data for evaluation were collected, therefore, from the remaining 34 schools.

[2] The private school met the eligibility criteria of offering sexual health education in middle school that is consistent with the state standards. As the curriculum was designed to be effective regardless of the type of schools, the study team decided to include this school in the sample.

The implementation period for Round 1 schools covered three semesters from Spring 2012 to Spring 2013; for Round 2 schools, two semesters from Fall 2012 to Spring 2013. The schools offered sexual health instruction as part of a semester-long health class, but not all schools offered a health class every semester (see Appendix C, Exhibit C.1). In most schools (26 of 34), sexual health education was offered only in 7th grade. In the remaining schools, sexual health education was offered in 7th and/or 8th grade.

All 7th and 8th grade students enrolled in a health class taught by a participating health education teacher during the implementation period were eligible to participate in the study if they had prior parental consent. The schools were assigned once, and students enrolled in the study classes were exposed to either Pono Choices or business-as-usual sexual health instruction according to their school's initial assignment. If a study teacher instructed more than two health classes in a semester, the first two classes in which the sexual health curriculum was taught were selected for the study.

Parental consent was collected for students enrolled in all eligible study classes, for both intervention and control schools, at the start of each semester. To eliminate potential influence of the assignment on parents' decision, we obtained their consent before they were informed about the curriculum used in their child's class. The consent form made no reference to the specific curriculum. Where teachers hosted a parent meeting to explain the curriculum, they did so after the due date for returning the forms. To encourage the timely collection of parental consent, students received a $10 gift card for returning a signed parental consent form regardless of whether

parents consented for them to participate in the study. In addition to parental consent, we obtained student assent at the time of baseline data collection.[3]

Overall, 2,208 students (1,383 in intervention and 820 in control) in the 34 participating schools were enrolled in study-eligible health classes during the study implementation period. Of those, we obtained parental consent for 1,783 students (1,158 in intervention group and 625 in control group).[4] Of those, 9 students were exposed to the intervention twice at one school that offered Pono Choices in 8th grade in Fall 2012 after offering it in 7th grade in Spring 2012. There was no crossover of study students across the conditions.

III.C.1 Data Collection for the Impact Evaluation

The intervention or business-as-usual curricula were implemented across study schools over three semesters from the spring of the 2011-12 school year through the spring of the 2012-13 school year. Depending on when the school was recruited and assigned and how often it offered a sexual health class, students from up to three semesters for each school were included in the study over the implementation period. Appendix C (Exhibits C.1 - C.3) provides a summary of the implementation and data collection timeline.

[3] The evaluation team collected assent forms from students whose parents had consented during administration of the baseline survey on the first day of the sexual health instruction. Before administering the survey, the evaluator explained the study to the students. Students were invited to ask questions about the study and the evaluation team answered questions, provided assurances of both the confidential and voluntary nature of the study, and asked the students to review and sign the assent form. Students declining to participate in the study (there were a total of seven across all student cohorts) were dismissed from the classroom to a pre-arranged location while the baseline survey was conducted. Students without parent consent were also dismissed from the classroom to the pre-arranged location. Students whose parents had consented to the study who declined to participate in the baseline survey were invited again to assent during administration of the follow-up survey.

[4] The study did not block schools by school size or classroom size, and the randomization resulted in the unintended difference in student sample sizes across the conditions. The sample size difference is attributed to several factors: (1) the total number of classes in the intervention group was 56, while total number of classes in the control group was 35 (8 schools in the intervention sample had 2 classes of students, compared to only 3 of the control schools having 2 classes; 5 intervention schools delivered sexual health instruction all three semesters, compared to only 1 of the control schools); (2) the average class size was larger in the intervention group, compared to the control group (21 vs. 18); and (3) the parent consent rate was lower in the control schools (76 percent) than in the intervention schools (84 percent) , even though parents were not informed of their school's treatment status until after the consent process was complete.

Impact analyses used two data sources: the student baseline survey conducted immediately prior to implementation of the sexual health curriculum and the follow-up survey conducted about 12 months after the baseline survey was administered. Survey instruments and procedures were identical for treatment and control schools. Of 1,783 students with parent consent, 1,735 took the baseline survey, and 1,548 took the 12-months post-baseline follow-up survey.

The evaluation team administered the baseline pre-curriculum surveys to students in classroom on the first day (or during the session immediately preceding the first day) of the sexual health instruction in both treatment and control schools, using scripted instructions in accordance with the survey administration protocols. The evaluation team administered the follow-up surveys to students who had parental consent, regardless of whether they assented to the baseline survey. The students were invited to participate in an onsite group survey administration session in their baseline middle school (for 7th graders at baseline) or at the high school they had moved to (for 8th graders at baseline). For students who were no longer in Hawai‘i middle or high schools or who missed the group administration, the evaluation team followed up individually and conducted the survey by mail, with phone and email follow-up. Students received a $10 gift card for completing the follow-up survey.

III.C.2 Data Collection for the Implementation Evaluation

The implementation study involved observing teacher training and classroom instruction, using structured instruments to measure fidelity and quality of implementation. The observers participated in a 2-day training that included observing videos of practice lessons conducted by study teachers. Each observer was paired with a senior “anchor” observer for at least two observations each semester, and refresher training was conducted at the beginning of each semester

to check for inter-rater reliability.[5] These practice observations and periodic paired observations throughout the study period were used to establish and maintain a 95 percent inter-rater reliability rate. Five Evaluation team members conducted classroom observations of 15 to 20 lessons each for a total of 128 lessons (22 percent of total lessons delivered over the three semesters), including at least two in each classroom each semester. Lessons to be observed were selected randomly, blocked by curriculum module such that each curriculum module was observed at least twice each semester.

The implementation team collected information on teaching experience from intervention schools teachers using a study enrollment form. The implementation team also interviewed control school teachers to learn their teaching background and determine the use of key components of Pono Choices in the counterfactual condition, and shared this information with the evaluation team. Appendix C (Exhibit C.3) summarizes data collection for the implementation study.

Fidelity. Teachers in intervention schools submitted facilitator logs documenting the activities covered during each lesson and modifications made, if any, to the lesson. The evaluation team collected observation logs for the 128 lessons observed, recording each lesson activity and the time spent on it.

Attendance (Dosage). Teachers submitted attendance data for all sexual health sessions they taught. Intervention teachers provided daily attendance logs and recorded the module and activities covered in each session. This information was used to calculate the percentage of the curriculum received by each student. Attendance logs were not collected in control schools.

Quality. During the classroom observations in intervention schools the evaluation team collected data on multiple aspects of program quality including the clarity of teachers' explanations,

[5] Inter-rater reliability was computed as the number of observation items where the raters concurred in their ratings divided by the total number of observation items on the instrument.

their enthusiasm, poise, and confidence, effectiveness of their response to students' questions, and level of student engagement.

Counterfactual. The implementation team conducted interviews with control school teachers after their final semester of participation. Information collected included key contents of the curriculum used and whether the curriculum was delivered by themselves or outside presenters.

Context. The implementation and evaluation teams collected limited contextual data through fidelity monitoring and ongoing contact with teachers and principals. No data were collected about students' exposures to sexual health services and education outside of school or after the delivery of the sexual health curriculum was completed.

III.D. Outcomes for Impact Analyses

Table 1 below summarizes the behavioral outcome measures used to address the primary and secondary research questions. The primary outcome measure was engagement in high-risk sexual behavior, which was based on survey questions about the use of condoms and birth control in the past three months. Students were regarded as engaging in high-risk sexual behavior if they reported having had sexual intercourse using neither condoms nor any effective means of birth control (Students were regarded as not engaging in high-risk behavior if they did not have sex or had sex using either condoms or other means of birth control.) For the secondary confirmatory question, the study examined the impact of the intervention on the initiation of sexual activity among youth. Students are regarded as having initiated sexual activity if they reported ever having sexual intercourse by the time of follow-up.

Table 1. Behavioral Outcomes Used for Primary and Secondary Research Questions

Outcome Name	Description of Outcome	Timing of Measure
Engagement in high-risk sexual behavior (outcome for primary impact research question)	The measure for engagement in high -risk sexual behavior was based on the questions on the use of condoms and birth control in the past 3 months: • "In the past 3 months have you had sexual intercourse without a condom?" • "In the past 3 months have you had sexual intercourse without an effective method of birth control – including condoms, birth control pills, the shot (Depo Provera), the patch, the ring (NuvaRing), IUD (Mirena or Paragard), implant (Implanon)?" Respondents are regarded as engaging in high-risk sexual behavior if they reported having used **neither** condoms nor birth control when they had sex.	1 year after baseline
Initiation of sexual activity (outcome for secondary impact research question)	The initiation of sexual activity among youth is measured by students' response ("Yes" or "No") to the first-year follow-up survey question: • "Have you ever had sexual intercourse?" If respondents did not respond to this question, but went on to answer subsequent questions and indicated they had sex, they are counted as having had sex.	1 year after baseline

In addition, the study explored the program impacts on non-behavioral outcomes, including students' knowledge of pregnancy and STI prevention, attitudes toward healthy sexual behavior, skills in managing relationships and choices, and intention to engage in safe-sex behavior. Appendix C (Exhibit C.4) summarizes the non-behavioral outcome measures used to examine the exploratory questions.

III.E. Study Sample

The target study sample is defined as all students who were enrolled in the target health education classes in the participating schools.[6] Of 2,203 students identified as eligible for the study, 1,783 received parental consent to participate. For each outcome analyzed, the analytic sample includes the students who completed survey items required to construct the measure on the follow-up survey. Of those 1,783 students with prior parental consent, 1,494 were included in the

[6] One exception was for a student who was unable to read, write, or understand English and for whom no translator was available. This student was considered ineligible for the study.

analytic sample for the estimation of impacts on engagement in high risk sexual behaviors, and 1,488 were included for the estimation of impact on the initiation of sexual activity.

Appendix C provides an overview of the sample sizes at key data collection points (Exhibits C.5.1-C.5.2) and provides a CONSORT diagram that summarizes the attrition and reasons for attrition in more detail (Exhibit C.6). A description of the study sample at baseline (for the sample of students completing baseline surveys) is also provided in Appendix C (Exhibit C.7).

III.F. Baseline Equivalence of Analytic Samples

To evaluate whether analytic samples were equivalent across assignment conditions, we compared select school and student baseline characteristics (including baseline outcome measures), of the treatment and control groups for each outcome analyzed (See Tables 2 and 3).

Table 2: Summary Statistics of Key Baseline Measures for Student Characteristics Analytic Sample for Estimating Impact on Engaging in High-Risk Sexual Behavior

Characteristic	Inter-vention N.	Inter-vention Mean	Inter-vention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value
Engagement in high-risk sexual behavior ("1"=used neither condom or birth control when having sex, "0"=otherwise)	843	0.01	0.08	453	0.00	0.05	0.00	0.160
Age in years	943	12.27	0.64	521	12.28	0.60	-0.01	0.589
Female ("1"=female, "0"=Otherwise)	949	0.52	0.50	525	0.54	0.50	-0.02	0.559
Student using non-English at home ("1"=Use non English at home; "0"=Otherwise)	961	0.27	0.44	533	0.39	0.49	-0.12**	0.000
Native Hawaiian ("1"=Native Hawaiian including mixed race, "0"=Otherwise)	948	0.41	0.49	529	0.49	0.50	-0.08	0.099
Asian ("1"=Asian including mixed race, "0"=Otherwise)	948	0.72	0.45	529	0.74	0.44	-0.02*	0.032

Characteristic	Inter-vention N.	Inter-vention Mean	Inter-vention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value
Non-Hispanic White ("1"= Non-Hispanic White including mixed race, "0"=Otherwise)	886	0.08	0.27	492	0.06	0.24	0.01	0.135
Hispanic ("1"=Hispanic, "0"=Otherwise	890	0.18	0.38	494	0.17	0.38	0.01	0.239

Source: IMPAQ staff calculations based on student surveys.
***Significant at the .01 level. * Significant at the .05 level. Notes: The total analytic sample size for this outcome was 1,494. Equivalence tests are reported based on the values of non-missing baseline characteristics. Race/ethnicity indicators are not mutually exclusive. See Appendix C Exhibit C.8 for mutually exclusive race/ethnicity categories and additional baseline characteristics tested for equivalence.*

Appendix C (Exhibits C.8 to C.16) summarizes additional details on equivalence tests for all impact analyses presented in this report, including exploratory analyses. The group difference for student characteristics was evaluated using the same statistical model used to estimate the program impact on outcomes (i.e., a mixed-level model in which each student-level baseline characteristic is regressed against the treatment status).[7] For confirmatory analytic samples, baseline characteristics examined were not statistically significantly different between the treatment and control groups, except for indicators for whether the student spoke English at home and whether the student identified themselves as Asian.[8]

[7] As noted below, the model was estimated using the dummy variable adjustment method for missing baseline characteristics. Equivalence tests are reported based on the values of non-missing baseline characteristic variables for the analytic sample.
[8] These two nonequivalent factors were included as covariates in the impact estimation.

Table 3. Summary Statistics of Key Baseline Measures for Student Characteristics Analytic Sample for Estimating Impact on Initiation of Sexual Activity

Characteristic	Inter-vention N.	Inter-vention Mean	Inter-vention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value
Initiation of sexual activity ("1"=Ever had sex, "0"=Otherwise)	825	0.06	0.25	440	0.07	0.26	-0.01	0.205
Age in years	941	12.27	0.64	518	12.28	0.60	-0.01	0.599
Female ("1"=Female, "0"=Otherwise)	946	0.53	0.50	522	0.55	0.50	-0.02	0.485
Student using non-English at home ("1"=Use non English at home; "0"=Otherwise)	958	0.27	0.44	530	0.39	0.49	-0.12**	0.000
Native Hawaiian ("1"=Native Hawaiian including mixed race, "0"=Otherwise)	945	0.41	0.49	526	0.49	0.50	-0.08	0.109
Asian ("1"=Asian including mixed race, "0"=Otherwise)	945	0.71	0.45	526	0.74	0.44	-0.03*	0.024
Non-Hispanic White ("1"=Non-HispanicWhite including mixed race, "0"=Otherwise)	883	0.08	0.27	490	0.06	0.24	0.01	0.130
Hispanic ("1"=Hispanic, "0"=Otherwise	887	0.18	0.38	492	0.17	0.38	0.01	0.251

Source: IMPAQ staff calculations based on student surveys.
***Significant at the .01 level. * Significant at the .05 level.*
Notes: The total analytic sample size for this outcome was 1,488. Equivalence tests are reported based on the values of non-missing baseline characteristics. Race/ethnicity indicators are not mutually exclusive. See Appendix C Exhibit C.9 for mutually exclusive race/ethnicity categories and additional baseline characteristics tested for equivalence.

III.G.1 Impact Evaluation Methods

For each confirmatory and exploratory outcome, program impacts were estimated as the differences in the student outcome measures between program and control groups, one year after baseline. The impact was estimated as intent-to-treat effects of the intervention, including all randomly assigned schools and students with consent to be in the study, regardless of the level of actual participation in the intervention. An intent-to-treat analysis maintains the initial randomized conditions, ensuring the estimated causal effects of the intervention on outcomes are unbiased.

The study specified a mixed-level model for impact estimation, in which students were nested within schools to account for the effects of clustering of individuals within each school. The model was applied to estimate an unbiased standard error of the impact estimates. The statistical model was used also to control for baseline outcome and sample characteristics, which improves the precision of the impact. Appendix E presents additional information on how estimation models were specified.

Selection of Covariates. Appendix C Exhibit C.21 presents a list of covariates considered and used for the results reported in this report. Besides study design controls (blocking and cohort), the estimated model included covariates found or expected to explain the variation in the outcome in the sample as well as a variable for which baseline equivalence was not established.[9] In selecting covariates, multicollinearity across covariates was also checked, and highly collinear variables were excluded.

Treatment of Missing Data. Not all students in the study sample took the surveys and not all of those who did take the surveys completed all survey items. Therefore, outcome and covariate data were missing for some students. The study applied listwise deletion to handle missing outcome data, assuming that data were missing at random. For missing covariate data, the dummy variable adjustment method was applied. For the dummy variable adjustment method, missing covariates were coded as zero for binary variables and as the sample mean for continuous variables. For these covariates, binary indicators for missing observations were included in the estimation.

[9] A set of covariates were selected based on the fit for the behavioral estimation models and on baseline equivalence testing for the behavioral outcome analysis samples. The same benchmark set of covariates were applied to the estimation of non-behavioral measures. For some non-behavioral analysis samples, an additional variable was found nonequivalent at baseline, but was not included in the reported estimation model. All covariates were included, however, in an alternative estimation model investigated in the sensitivity analyses.

Sensitivity Analyses. To assess whether impact estimation results would remain robust across a range of methodological choices, we conducted several sensitivity analyses including:

- ***Alternative methods for treating missing covariates:*** As an alternative to the dummy variable adjustment method, the evaluation team applied a listwise deletion method to missing covariates.
- ***Alternative covariate specification:*** Alternative sets of student- and school-level baseline variables were applied. (See Appendix C Exhibit C.21 for a list of the alternative sets.)
- ***Alternative estimation model:*** For the confirmatory analyses, the evaluation team applied the maximum likelihood method to estimate a random intercept logit model. As an alternative approach, a random intercept probit model is estimated.
- ***Alternative sample:*** To check for and minimize bias due to the attrition by two schools that dropped out prior to implementation, the study conducted impact analyses using an alternative sample based on 31 schools by excluding the remaining schools in the blocks affected by the attrition.

III.G.2 Implementation Evaluation Methods

Appendix C (Exhibit C.22) summarizes the implementation evaluation methods used.

Fidelity. Content delivered was measured as the percentage of intended activities completed in each program module as reported by both teachers and observers. The changes to the Pono Choices curriculum documented in facilitator (teacher) and observer logs were reviewed by the evaluation, implementation, and curriculum development teams each semester to identify if there were substantial adaptations. The evaluation team also compared teachers' and observers' assessments of the activities completed for the 128 sessions observed.

Dosage. Dosage was calculated as the percentage of modules completed. In cases where modules were not completed within a single session, students were counted as having completed the module if they completed at least 80 percent of the activities in the module. Percentage attended was calculated as the number of modules attended divided by a total of 10 modules.

Quality. Quality of the delivery of instruction was measured using 11 items on the Program Observation Form, rated on a 5-point scale, where 5 is excellent. These observer ratings were summarized as average scores and as a percentage of lessons rated with an average score across the 11 observation items of at least 4.0 on the 5-point scale. One of the 11 items observed was of particular interest: Quality of student engagement was calculated as 1) the percentage of the student engagement observation items rated by the evaluator as at least 4.0 on the 5-point scale and 2) the average student engagement score across teachers and modules.

Counterfactual. The implementation team summarized data from interviews with control teachers and provided the evaluators with a spreadsheet. Data are presented as frequency counts and percentages.

Context. The context for provision of TPP instruction was documented by implementation staff and evaluators as part of fidelity monitoring and maintenance. Any contextual factors noted by team members or arising in the local media were discussed at monthly meetings of implementation and evaluation team members to assess potential effects on implementation.

IV. STUDY FINDINGS

IV.A. Implementation Study Findings

The study examined fidelity, dosage, and quality of the intervention as implemented. Appendix F (Exhibit F.1) provides a summary of implementation analysis findings.

Fidelity. Teachers completed 98 percent of intended activities across all three semesters of implementation. A comparison of teachers' reports vs. observer assessment of 128 lessons observed by the evaluation team showed 98 percent concurrence. (See Appendix F Exhibit F.2 – F.4). As shown in Appendix F Exhibit F.6, the time spent on each module was longer during the first semester of implementation, with the average time spent on each module exceeding the intended 60 minutes for 7 of the 10 modules (see Exhibit F.6.1). The average time spent per module during the second and third semesters of implementation was actually slightly less than intended across all 10 modules See Exhibit F.6.2). This reflects an increased emphasis on lesson pacing during the Round 2 teacher training and the refresher trainings for the second and third semesters, as well as the fact that some of the teachers in the second and third semesters had by then gained experience delivering the curriculum during the first semester. The longer-than-intended delivery time in the first semester and among teachers who continued to spend more time than intended in later semesters was determined through observations to be due to variations in teaching and classroom management style rather than to deviation from the intended curriculum.

Dosage. Ninety-four percent of students completed at least 75 percent of the curriculum, and the average dosage (percent of modules attended) was 94 percent across all three semester cohorts. Attendance varied slightly across different program modules, ranging from about 92 percent to almost 96 percent. (See Appendix F Exhibit F.5.)

Quality. As shown in Appendix F Exhibit F.1, for all three semesters of implementation, the average overall rating across the 11 measures of quality of the delivery of the curriculum was 4.27 on a scale of 1-5, where 5 is Excellent. As shown in Exhibit F.1, each semester the percentage of modules with an average score of 4.0 or higher ranged from 73 percent to 83 percent. Exhibit F.7 shows the ratings by module.)

Student Engagement. Among the classroom observation items of program quality, one area of particular interest to the project was student engagement. As shown in Appendix F Exhibit F.1, the average student engagement rating was 4.61 on a scale of 1-5, where 5 is Excellent. Exhibit F.1 shows that across all three semesters, the percentage of lessons with a score of 4.0 or higher was 86 percent. (Exhibit F.8 shows the ratings by module.)

Counterfactual. Some control school teachers reported providing only one or two sessions of sexual health curriculum, while others provided up to 10 sessions. While the Hawai'i Department of Education identified specific approved curricula that schools could choose from, schools were free to choose how much of the material to use. Only one of the control schools fully implemented an approved sexual health curriculum, using *Making a Difference* one semester and *Making Proud Choices* another semester. (A total of 4 percent of all control student received one of these evidence-based programs.) Other schools sometimes relied on outside presenters (affecting about 40 percent of control students) who provided their own curriculum (including the AIDS foundation, YMCA, Maui Youth and Family Services, Bay Clinic). Some of the public charter schools focused on project-based learning, focusing on any sexual health topics that students chose for school projects rather than offering direct instruction.

Five key components of the Pono Choices curriculum—reproductive anatomy, pregnancy prevention, STI prevention, refusal skills, and condom demonstration—were included in the business-as-usual program in at least some of the control schools. As shown in Appendix F Exhibit F.10, four of these key curriculum components (reproductive anatomy, pregnancy prevention, STI prevention, and refusal skills) were present in the majority of control schools involved in the study and received by the majority of control students (between 64.3 and 87.7 percent). The fifth key component, the condom demonstration, was included at 5 of the 17 control schools, reaching 23.7

percent of control students. Only 2 of the 17 control schools included all five components (delivered to 9 percent of total sample of control students).

As shown in Exhibit F.11, 37.5 percent of the control schools had a teacher deliver the entire curriculum, and one-half brought in outside experts to assist the teacher. In two control schools, only outside experts taught the sexual health class. This situation differed from the program schools where regular classroom teachers delivered 100 percent of the curriculum. Exhibit F.11 shows similar levels of teaching experience between program and control teachers.

IV.B. Confirmatory Impact Analysis Findings

The impact analysis of behavioral measures found that participation in Pono Choices had no detectable behavioral impacts on the analytic sample of students one year after baseline. Exhibit G.1 in Appendix G provides unadjusted summary statistics of the behavioral outcome measures. Table 4 below reports the (regression adjusted) estimated effects of the intervention for the primary and secondary behavioral outcomes. Exhibits G.3-G.4 in Appendix G provide additional detail on the regression-adjusted estimated impacts of Pono Choices in terms of estimated probabilities of behavioral outcomes. The impact estimates in the appendix are also reported in terms of an odds ratio (relative risk). For example, the odds ratio for unsafe sex is the ratio of the probability of engaging in unsafe sex to the probability of not engaging in high-risk sexual behavior.

Table 4. Post-Intervention Estimated Effects Using Data from Student Survey to Address Primary and Secondary Research Questions

Outcome Measure	Intervention	Comparison	Treatment Effect (*p*-value of difference)
Engagement in high-risk sexual behavior, one year after baseline (1=used neither condom or birth control during intercourse, 0=did not have intercourse, or used either condom or birth control, or both, during intercourse)	0.015	0.022	-0.007 (0.428)
Initiation of sexual activity, one year after baseline (1=ever had sex by the follow-up, 0=otherwise)	0.098	0.100	-0.001 (0.944)

Source: IMPAQ staff estimation based on student surveys.

Impact on Engagement in High-Risk Sexual Behavior (Primary Question). The estimated probability (or the estimated percent) of engaging in high-risk sexual behaviors among the treatment group was 1.5 percent, while it was 2.2 percent among the control group (see Table 4). However, the difference—the estimated impact of Pono Choices—was not statistically significant.

Impact on Initiation of Sexual Activity (Secondary Question). The estimated probability (or estimated percent) of ever having had sex at first year follow-up for the treatment group was 9.8 percent, while it was 10.0 percent for the control group (See Table 4 above). The difference was not statistically significant.

Sensitivity Analyses. All sensitivity analyses conducted yielded consistent results, showing no statistically significant effects on engagement in high-risk sexual behavior or initiation of sexual activity. Exhibits G.6–G.7 report the results of key sensitivity analyses of the behavioral outcomes.

IV.C. Exploratory Impact Analysis Findings

Exhibit G.1 in Appendix G reports the unadjusted summary statistics of non-behavioral outcome measures at one year after baseline. Exhibit G.5 reports the estimated impacts on non-behavioral outcomes.

The exploratory analysis found that participation in Pono Choices had a detectable impact on student knowledge of pregnancy and STI prevention at one year after baseline.[10] On average, those in the treatment group answered 71.8 percent of knowledge questions correctly, while those in the control group answered 56.4 percent correctly. The difference between the groups (15.4 percentage points) is statistically significant at the 0.01 level ($p < 0.001$), and the effect size is 0.788. Sensitivity analyses conducted on the knowledge measures show that these results are large and robust across alternative technical specifications (See Exhibit G.8). No significant effects were found, however, on other non-behavioral measures: attitudes toward healthy sexual behavior, skills in managing relationships and choices, and intention to engage in safe-sex behavior skills.

V. CONCLUSIONS

V.A. Summary and Implications of Findings

Impact analyses found no effect of the program on youth behavior, but a significant impact on student knowledge. There were no statistically significant impacts on either engagement in high-risk sexual behaviors or initiation of sexual activity within the 1-year observation period. Students in the treatment group were no more or less likely to report having engaged in high-risk

[10] The pregnancy/STI prevention knowledge analyzed in the study included knowledge about condom use as one of its components. The study also examined the condom use knowledge separately from the pregnancy/STI prevention knowledge as well as pregnancy/STI knowledge without condom use knowledge as its component. The results based on all these measures were similar. This reports presents the result form the overall knowledge measure which is aligned with the broader goal of the intervention.

sexual behaviors in the previous three months than the control group. Similarly, they were no more or less likely to report the initiation of sexual activity at follow-up than the control group one year after baseline.

The failure to detect statistically significant impacts in these behavioral outcomes was not surprising given the relatively young age of the students at baseline (average age of 12). As noted earlier, the study was originally designed to assess the impact of Pono Choices two years after baseline, when more students reach the age to start encountering occasions to apply the knowledge and skills gained through Pono Choices, but the collection for the second year follow-up data was unexpectedly not allowable. The results at the 1-year follow up may not be fully capturing the intended impacts of the intervention, since the majority of the students were yet to become sexually active. In our sample, an unadjusted percentage of control group students having ever had sex increased by just 2 percentage points from 9 at baseline to 11 percent at one year after baseline, underscoring the challenge of detecting a program impact that is likely to be small. A longer-term follow-up is warranted to investigate the potential impacts on behavioral outcomes.

Two characteristics of business-as-usual curriculum delivery in control schools might also have affected the detection of impacts on student behavior: (1) the majority (between 64.3 and 87.7 percent) of control group students received key curriculum components of Pono Choices curriculum, and (2) 35 percent of the control youth were taught by trained outside experts (versus none of treatment group). Both conditions could have narrowed the effective contrast between the intervention and control groups.

Exploratory analyses found that the treatment group students scored statistically significantly higher than the control group on knowledge about pregnancy and STI prevention at one year after baseline. The knowledge measures assessed the student comprehension of topics

such as methods to prevent pregnancy and STI, proper use of condoms, and types of birth control (see Appendix C Exhibit C.4 for more details on the measure). The findings suggest that an effect of Pono Choices on knowledge could potentially lead to changes in behavior, pointing to the need for further investigation of the long-term effects of Pono Choices.

The failure to detect effects on other non-behavioral measures—skills, attitudes, intention—one year after baseline suggests several possibilities to investigate. For example, Pono Choices' influence could be limited to students' knowledge and understanding of pregnancy and STI prevention. Alternatively, the findings could mean that attitudes, skills, and intention were not as proximal as they were originally assumed. Yet another consideration is measurement limitations. These non-behavioral measures—the average of 4-point scale ratings—may not have been adequate to capture the group difference. The unadjusted control group means of the average ratings on attitudes, skills, and intentions at follow-up ranged from 3.0 to 3.7, with 4.0 being the highest, leaving little opportunity for a treatment effect.

V.B. Study Limitations

The study has important limitations, including the limited generalizability of findings, limitation with the randomized design, and sample attrition.

Limited ability to generalize results. The schools were purposively selected based on their willingness and administrative support of schools' participation in the study. Given the intentional selection process, the findings from this study are not generalizable to all Hawai'i schools or students.

Limited ability to assess the importance of cultural responsiveness. The cultural responsiveness of the Pono Choices curriculum was considered a high priority and key focus of curriculum development. However this study was designed to test the Pono Choices curriculum against business as usual and was not designed to test the cultural aspects of the curriculum. While the intervention group students scored statistically significantly higher at follow-up than the control group on knowledge about pregnancy and STI prevention, without contrasting the contents of the Pono Choices with and without the cultural components, it is not possible to assess how much of the difference in knowledge gains might be due to the cultural aspects of the curriculum.

Sample attrition. The study faced minimal attrition at the school level (2 of 36 schools dropped out after random assignment). At the student level, response rates for completing the surveys were high (a total of 86.8 percent of students whose parents consented to the study completed the 1-year follow up – 86.1 percent among intervention students and 88.2 percent among control students). However, parent consent rates differed between intervention schools and controls, which could lead to bias in the impact estimates: the parent consent rate was 83.5 percent for intervention school students and 73.2 percent for controls. (See Appendix C Exhibit C.5.1 for student attrition for each outcome measure.) To address the potential bias due to small amount of

school attrition, we conducted sensitivity analyses by excluding the blocks that included the two withdrawn schools and confirmed that findings based on the alternative and original samples were consistent.[11] Similarly, we checked for baseline equivalence of observable characteristics of students in the analytic samples and controlled them as covariates in estimating impacts. However, potential bias due to the school and student attrition may remain.

Limitation on the application of the random assignment design. The original purpose of the study was to examine the intent-to-treat effects on individuals exposed to the intervention. Like many cluster randomized controlled trials conducted in school settings, randomization for this study had to be done before study-eligible students could be identified. To maintain the randomized conditions for individual-level inferences, the study made the assumptions that: (a) enrollment in study schools was independent of the assignment; and (b) parent consent to participate was obtained without knowledge of the assignment status. The study then regarded the resulting student sample as equivalent to a sample drawn at the time of random assignment. These assumptions, however, are not verifiable.

[11] The sample and analysis information for the alternative 31 school sample is provided in the appendix: See Appendix C Exhibit C.6 for the sample flow chart; see Exhibit C.17-C.19 for the baseline characteristics; see Appendix G Exhibit G.2 for a summary of the behavioral measures; see Exhibits G.6-G.8 for the estimation results for the behavioral and knowledge measures.

VI. REFERENCES

Centers for Disease Control and Prevention. 2011 Sexually Transmitted Diseases Surveillance, Chlamydia Profiles. http://www.cdc.gov/std/chlamydia2011/. Published 2015.

Centers for Disease Control and Prevention. (2012a). 1991-2011 High School Youth Risk Behavior Survey Data. http://www.hhs.gov/ash/oah/adolescent-health-topics/reproductive-health/states/hi.html.

Centers for Disease Control and Prevention. (2012b) Hawai'i 2013 State Health Profile. http://www.cdc.gov/nchhstp/stateprofiles/pdf/Hawaii_profile.pdf

Centers for Disease Control and Prevention. (2015) Teen Pregnancy. http://www.cdc.gov/teenpregnancy/about/index.htm.

Eaton DK, Kann L, Kinchen S, et al. Youth Risk Behavior Surveillance — United States, 2011. *Morbidity and Mortality Weekly Report*. 2012;61(SS04):1-162.

Frost JJ, Lindberg LD, Finer LB. Young Adults' Contraceptive Knowledge, Norms and Attitudes: Associations with Risk Of Unintended Pregnancy. *Perspectives on Sexual and Reproductive Health*. 2012;44(2):107-116. doi:10.1363/4410712.

Hamilton BE, Ventura SJ, Mathews TJ. National and State Patterns of Teen Births in the United States, 1940–2013. *National Center for Health Statistics*. 2014;63(4). http://www.cdc.gov/nchs/data/nvsr/nvsr63/nvsr63_04.pdf.

Hawaii Department of Education. Hawaii Content & Performance Standards III Database. http://165.248.72.55/hcpsv3/. Published June 2007.

Hawaii Department of Education. Sexual activity, Pregnancy and STIs Among Hawaii Teens, 2011 Youth Risk Behavior Survey. Sexual Health Education. http://www.hawaiipublicschools.org/TeachingAndLearning/HealthAndNutrition/sexed/Pages/default.aspx. Published October 24, 2015. Accessed October 24, 2015.

Hawkins JD, Catalano RF, Kosterman R, Abbott R, Hill KG. Preventing adolescent health-risk behaviors by strengthening protection during childhood. *Arch Pediatr Adolesc Med*. 1999;153(3):226-234.

Kornides ML, Kitsantas P, Lindley LL, Wu H. Factors Associated with Young Adults' Pregnancy Likelihood. *Journal of Midwifery & Women's Health*. 2015;60(2):158-168. doi:10.1111/jmwh.12258.

Kost, Henshaw. U.S. Teenage Pregnancies, Births and Abortions, 2010: National and State Trends by Age, Race and Ethnicity. *Guttmarcher Institute*. May 2014. http://www.guttmacher.org/pubs/USTPtrends10.pdf.

Ladson-Billings G. But that's just good teaching! The case for culturally relevant pedagogy. *Theory Into Practice*. 1995;34(3):159-165. doi:10.1080/00405849509543675.

Moore KA, Sugland BW. Next Steps and Best Bets: Approaches to Preventing Adolescent Childbearing. *Child Trends*. February 2001. http://eric.ed.gov/?id=ED414555. Accessed August 27, 2015.

State of Hawai'i, Board of Education. Abstinence-Based Education Policy (2010). Adopted 1995. http://www.hawaiiboe.net/policies/2100series/Pages/2110.aspx. Accessed October 20, 2014.

U.S. Census Bureau. 2010 Census. October 2015. http://factfinder.census.gov. Accessed October 24, 2015.

U.S. Department of Health and Human Services. Hawaii Adolescent Reproductive Health Facts. Office of Adolescent Health. http://www.hhs.gov/ash/oah/. Updated November 13, 2014.

APPENDIX A. PONO CHOICES CURRICULUM DESCRIPTION

The Pono Choices curriculum utilizes a place-based approach, whereby the local traditions, Hawaiian cultural practices, and specific place, engage and provide access for local students' introduction to sexual health topics, and reinforce the teen pregnancy and STI prevention message. Place-based education differs from conventional text and classroom-based education in that it draws from students' local community as one of the primary resources for learning. Place-based education promotes learning that is rooted in what is local—the unique history, environment, culture, economy, literature, and art of a particular place—that is, in students' own "place" or immediate schoolyard, neighborhood, town or community (Manaseri, Roberts, Stofocik, Manuel & Uehara, 2014).

Pono Choices draws from social learning, self-regulation, and developmental assets theories. Social learning theory (Brindis, Sattley, & Mamo, 2005) focuses on behavior as a result of continuing interaction between a person, the behavior of that person, and the environment within which the behavior is performed. Major concepts include skill building, including goal-directedness, emotional coping, and problem solving. The second theory incorporated into the Pono Choices curriculum is self-regulation. The premise of this theory is that individuals operate like feedback systems, constantly regulating their relationships to the environment in order to bring their current states closer to their goal states. This theory emphasizes coping procedures and problem solving. Finally, the Pono Choices curriculum utilizes the developmental assets/resiliency model. This theory seeks to enable youth to participate in socially useful tasks so that they become healthy adults, in spite of adversity, and demonstrate positive results in self-esteem and moral development (Lerner & Benson, 2003).

Pono Choices was developed through a collaborative workgroup process. The curriculum development process began with creating community partnerships specifically with Planned Parenthood of Hawai'i (PPHI) and ALU LIKE, Inc. (ALI), a nonprofit native Hawaiian serving organization. These partnerships were then nurtured and maintained throughout the development process. PPHI and ALI provided expertise in ensuring that both medically accurate sexual health terminology and culturally responsive concepts respectively, were embedded throughout the lessons. The University of Hawai'i also consulted with the developers of an award-winning evidence-based teen pregnancy prevention curriculum, and used a participatory process involving students, teachers, members of the Hawai'i Department of Education and other community stakeholders to inform the Pono Choices curriculum. Once a core writing team was identified, the creation of Pono Choices took place in two major stages: (1) setting the foundation which included writing the content and (2) pilot testing (Manaseri, Uehara & Roberts, 2013).

Knowledge, attitudes, and skills about pregnancy and STI prevention are reinforced in the curriculum through cultural referents in four essential activities: 1) an introduction of a Hawaiian cultural value at the beginning of each module, 2) an original cultural story in 10 parts, one to introduce each lesson, 3) cultural practices that are shared as take-home activities in select modules, and 4) locally produced videos that present the topic or intended message.

Hawaiian cultural values are introduced in a Hawaiian language term in each lesson, like a word of the day, to reinforce lesson content and are expanded upon through the original cultural story entitled "The Voyage of the Wa'a Kaulua." This audio story, paired with original artwork, was created specifically for the curriculum and serves as an access point into the curriculum content from the viewpoint of two adolescents preparing for an important journey. It features two youth, Ka'iwi and Pailolo, who are going through puberty and journeying towards their goals and

dreams as they transition to adulthood. Through "The Voyage of the Wa'a Kaulua," important values are introduced through the Hawaiian cultural practice of oral history and story-telling. Cultural practices shape thinking processes, which serve as tools for learning within and outside of school (Hollins, 1996). A 10-minute audio recording of a chapter of the cultural story is used to introduce the Hawaiian cultural value that is associated with each lesson.

In addition to the above, Pono Choices embeds cultural practices in the curriculum through 'ohana (family) activities. Students and members of their 'ohana have the opportunity to construct a wa'a (canoe), braid cordage, and create a lei while reinforcing the messages of teen pregnancy and STI prevention. These 'ohana activities serve as an opportunity to bridge what students are learning in school to a family's experience, values, and beliefs about this sensitive subject area. Along with the stories, Pono Choices also uses locally produced videos and historical readings throughout the curriculum to connect students to their community. The following modules comprise the 9.5 hour program.

Module 1: An Introduction to Pono Choices

- Purpose of teen pregnancy and sexually transmitted infection prevention education
- Pono Choices cultural representation of the waʻa (Hawaiian voyaging canoe) and how it relates to sexual health

Module 2: Pono—Making Pono Choices

- Identification of key people who can help students make pono choices
- Definitions of sex and abstinence
- Media messages about sex
- Short-term and long-term goal setting

Module 3: Mōhala—Lessons in Anatomy and Puberty

- The parts and functions of the male and female reproductive systems
- Body changes during puberty

Module 4: Nohona—The Role of Communication in Healthy Relationships

- Elements of healthy, unhealthy, and abusive relationships
- Support system for reaching goals
- The role of communication in developing and maintaining healthy relationships

Module 5: Aloha—Maintaining Respect in Relationships

- Refusal skills
- Effective use of refusal skills in pressure situations
- Alternative ways to show affection other than sexual intercourse

Module 6: Hāpai Pono—Pregnancy

- Emotional, physical, and financial responsibilities for pregnancy
- Financial aspects of child rearing
- Methods of birth control, including abstinence, hormonal and barrier methods

Module 7: Pilina A‘o—Understanding Sexually Transmitted Infections

- Types of STIs and how they are transmitted
- Bodily fluids that transmit HIV and other STIs
- Myths surrounding STIs

Module 8: Pilina Pono—Preventing Sexually Transmitted Infections

- STI prevention
- Effective condom use
- Risk factors of different sexual behaviors

Module 9: Nā Kūlia—Negotiation Skills

- Negotiation and decision making skills
- Refusing unsafe or unwanted behavior
- Refusal skills practice

Module 10: Oli Ho‘omana—Empowerment

- Review of various teen pregnancy and STI prevention methods
- Review of students’ immediate and long-term goals
- Review of key vocabulary and concepts.

References for Appendix A

Brindis, C.D., Sattley, D., Mamo, L. (2005). From theory to action: Frameworks for implementing community-wide adolescent pregnancy prevention strategies. San Francisco, CA: University of California, San Francisco, Bixby Center for Reproductive Health Research and Policy, Department of Obstetrics, Gynecology and Reproductive Sciences, and the Institute for Health Policy Studies.

Hollins, E. R. (1996). Culture in school learning. New Jersey: Lawrence Erlbaum Associates Inc.

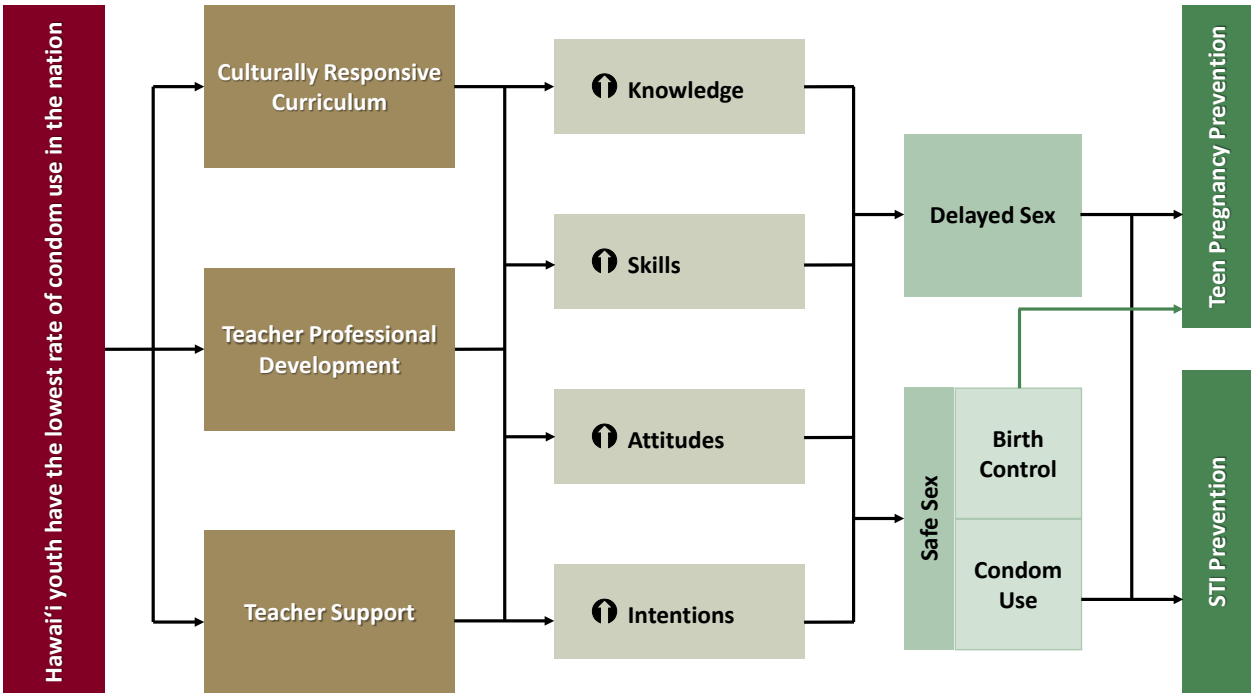
Lerner, R. M., & Benson, P. L. (2003). Developmental assets and asset-building communities: Implications for research, policy, and practice. New York, NY: Kluwer Academic/Plenum Publishers.

Manaseri, H., Uehara, D., Roberts, K. (2013). *Making Pono Choices: A Collaborative Approach to a Culturally Responsive Teen Pregnancy and Sexually Transmitted Infection Curriculum in Hawai'i . Maternal and Child Health Journal*, Volume 17, No. 10.

Manaseri, H., Roberts, K., Stofocik, K., Manuel, N., Uehara, D. (2014) Culture as a Protective Factor: The Use of Storytelling in a Teen Pregnancy and STI Prevention Curriculum. *Health Care Current Reviews* 2: 127. doi:dx.doi.org/10.4172

APPENDIX B. LOGIC MODEL OF INTERVENTION AND OUTCOMES

Exhibit B.1: Pono Choices Logic Model



APPENDIX C. SUPPLEMENTAL EXHIBITS FOR CHAPTER III

Exhibit C.1: Implementation Schedule—School Participation by Semester

Schools	SP2012	F2012	SP2013
First Round Intervention Schools			
Intervention School 1.1	2-7-2012	8-16-2012	
Intervention School 1.2	4-23-2012	11-13-2012	4-17-2013
Intervention School 1.3	4-11-2012	10-22-2012	3-28-2013
Intervention School 1.4	1-23-2012		1-25-2013
Intervention School 1.5	1-25-2012	10-11-2012	4-11-2013
Intervention School 1.6	4-2-2012	10-29-2012	4-3-2013
Intervention School 1.7	4-23-2012	11-16-2012	4-22-2013
Intervention School 1.8	2-27-2012		1-25-2013
Intervention School 1.9	2-21-2012	10-10-2012	
First Round Control Schools			
Control School 1.1		9-19-2012	3-13-2013
Control School 1.2	5-16-2012	10-24-2012	
Control School 1.3	1-4-2012	8-27-2012	
Control School 1.4	2-14-2012	11-1-2012	2-7-2013
Control School 1.5	2-23-2012	11-1-2012	
Control School 1.6	3-19-2012		4-3-2013
Control School 1.7	2-27-2012		12-4-2012
Control School 1.8	3-30-2012		4-3-2013
Control School 1.9	5-15-2012	9-20-2012	2-21-2013
Control School 1.10			
Second Round Intervention Schools			
Intervention School 2.1		9-26-2012	3-6-2013
Intervention School 2.2		10-15-2012	4-24-2013
Intervention School 2.3		10-09-2102	.
Intervention School 2.4		11-8-2012	4-15-2013
Intervention School 2.5		8-28-2012	
Intervention School 2.6		8-24-2012	
Intervention School 2.7		10-11-2012	4-5-2013
Intervention School 2.8		8-27-2012	1-15-2013
Second Round Control Schools			
Control School 2.1			2-25-2013
Control School 2.2		10-15-2012	
Control School 2.3		9-17-2012	
Control School 2.4			1-14-2013
Control School 2.5			2-11-2013
Control School 2.6		11-5-2012	1-11-2013
Control School 2.7			3-5-2013

Exhibit C.2: Implementation and Data Collection Timelines

Data Collection Effort	SP2012	F2012	SP2013
Start date of intervention[a]	01/05/12	08/23/12	01/15/13
Impact Evaluation Data			
Pre-curriculum baseline survey	01/04/12 – 05/16/12	08/16/12 – 11/16/12	12/04/12[b] – 04/22/13
1-year follow-up survey	01/08/13 – 04/18/13	08/21/13 – 11/22/13	12/02/13 – 05/09/14

(a) Earliest date of curriculum implementation. Actual start date varies by school.
(b) One school requested the evaluators collect baseline data at the end of fall 2012, although instruction actually began in January 2013.

Exhibit C.3: Implementation Data Collection Summary

Implementation Element	Types of Data Used to Assess whether the Element of the Intervention Was Implemented as Intended	Frequency/Sampling of Data Collection	Party Responsible for Data Collection
Adherence			
How often were sessions offered? How many were offered?	For each module of the curriculum, facilitator logs were used to collect data on: • The number of activities planned • The number of activities actually delivered • Whether each activity was delivered as intended or with changes and what the changes were	Facilitator logs were completed for every session and submitted weekly	Teachers delivering the curriculum (submitted to the external evaluator)
	For a sample of modules observed by evaluators, observation logs were used to collect data on: • The number of activities planned • The number of activities actually delivered • Whether each activity was delivered as intended or with changes, and what the changes were • The number of minutes planned vs. actual for each activity	Observations were conducted at least three times per semester in each participating class in each intervention school.	External evaluation staff
What and how much was received?	Attendance logs were used to collect individual student attendance data for all sessions.	Attendance data were submitted weekly for each participating class in each intervention school.	Teachers delivering the curriculum (submitted to the external evaluator)

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
What content was delivered to youth?	For each module of the curriculum, facilitator logs were used to collect data on: • The curriculum components planned • The curriculum components actually delivered • Whether each component was delivered as intended or with changes, and what the changes were For a sample of modules observed by evaluators, observation logs were used to collect data on: • The number of activities planned • The number of activities actually delivered • Whether each activity was delivered as intended or with changes, and what the changes were	Facilitator logs were completed for every session and submitted weekly. Observations were conducted at least three times per semester for each participating class in each intervention school.	Teachers delivering the curriculum (submitted to the external evaluator) External evaluation staff
Who delivered material to youth?	List of teachers trained to implement the curriculum and receiving ongoing support with number of years of teaching experience	Information collected at enrollment for all participating teachers.	Program Implementation staff
Quality			
Quality of staff-participant interactions	Program Observation Forms were used to collect data on quality of implementation of the program delivery. The form used was provided by OAH and included ratings of the clarity of teachers' explanations, their enthusiasm, poise and confidence, how effectively they responded to students' questions, and level of youth student engagement with the lesson on a scale of "1" to "5" where 5 was the highest possible rating for each item.	Observations were conducted at least three times per semester in each participating class in each intervention school	External evaluation staff

Implementation Element	Types of Data Used to Assess whether the Element of the Intervention Was Implemented as Intended	Frequency/Sampling of Data Collection	Party Responsible for Data Collection
Quality of youth engagement with program	Program Observation Forms were used to collect data on quality implementation of the program delivery. The form used was provided by OAH and included ratings of the clarity of teachers' explanations, their enthusiasm, poise and confidence, how effectively they responded to students' questions, and level of youth student engagement with the lesson on a scale of "1" to "5" where 5 was the highest possible rating for each item. Item #5 rated "How actively did students participate in discussions and activities?"	Observations were conducted at least three times per semester in each participating class in each intervention school	External evaluation staff
Counterfactual			
Experiences of counterfactual condition	Exit interviews with control teachers: • Teaching experience (sexual health and total years) • Facilitator type (teacher vs. outside presenter) • Content areas covered include: ○ Reproductive Anatomy ○ Pregnancy prevention ○ STI prevention ○ Refusal skills ○ Condom demonstrations	Interviews were conducted with all control school teachers at the end of the study period.	Program Implementation staff
Context			
Other TPP programming available or offered to study participants (both intervention and counterfactual)	Service provision was monitored by implementation staff and evaluators as part of the fidelity monitoring and maintenance during implementation. Information about the content and duration of the control school sexual health curriculum was collected during exit interviews at the end of each semester they participated to describe the counterfactual condition. No data were collected about other exposures to sexual health information, either outside of school, or after the delivery of sexual health curriculum was completed, nor do we have access to any other source of this information	Ongoing and at the end of each semester	Implementation and External Evaluation Teams
External events affecting implementation	Monitored news media; checked in with participating teachers periodically	At least monthly	Implementation and External Evaluation Teams

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
Substantial unplanned adaptation(s)	Unplanned changes or adaptations during delivery of the Pono Choices curriculum were captured on facilitator logs and observer fidelity logs.	Facilitators' logs were completed for every session and submitted weekly. Observations were conducted at least three times per semester in each participating class in each intervention school.	Teachers delivering the curriculum External evaluation staff

Exhibit C.4: Non-behavioral Outcomes Measures for Exploratory Analyses

Outcome Name	Description of Outcome	Timing of Measure Relative to Program
Knowledge of pregnancy and STI prevention	This measure is composed using a weighted average, giving 90% weight to the percentage of items that respondents answered correctly regarding knowledge of pregnancy and STI prevention and 10% weight to the percentage correctly answered on questions about knowledge of condom use. If respondents answered some items but skipped others, only valid, non-missing responses are counted in the calculations. When respondents skipped all questions in the section, their score is coded as missing. The 9 multiple-choice items for knowledge of pregnancy and STI prevention covers the following: 1) Effective way to prevent pregnancy 2) Behaviors associated with high risk of HIV 3) Definition of viral STI 4) Identification of STI carrier 5) Pregnancy 6) Prevention methods for STI 7) Refusal skills 8) Transmission mechanism of HIV 9) Types of birth control. (For the baseline survey, it includes the 10th item on condom use.) The 9 multiple choice items for condom use in the follow-up survey covers the following: 1) Condom usage 2) Wearing two condoms 3) How far unrolled 4) No space at tip of condom 5) Rolling wrong way 6) Expiration date 7) Removal of condom 8) Lubricants 9) Wallet for storage	1 year after baseline

Outcome Name	Description of Outcome	Timing of Measure Relative to Program
Attitudes toward healthy sexual behavior	This measure is an average rating of importance of 10 healthy sexual behaviors. The measure is calculated from 10 items on the survey. Participants were asked how important the following are: • Not having sex until I am ready • Avoiding risky sexual behavior • Preventing unwanted pregnancy as a teenager • Knowing what kind of birth control methods I can use to prevent an unwanted pregnancy • Taking personal responsibility for my sexual health • Communicating openly about sexual intent with my partner • Knowing multiple ways to prevent STIs and unwanted pregnancy • Using condoms to prevent STIs and unwanted pregnancy • Using alternative ways to show affection other than having sex • Understanding changes that happen during puberty The variable is constructed as a continuous variable by taking the average of items validly answered. Values range from "0" (Not Important) to "4" (Very Important). If respondents answered some items but skipped others, only valid, non-missing responses are counted in the calculations. When respondents skipped all questions in the section, their score is coded as missing. The composite score based on 10 items has an alpha of 0.86 at baseline and 0.88 at 1-year follow-up.	1 year after baseline
Skills in managing relationships and choices	This measure is an average rating of difficulty of five skills related to managing relationships and choices. The measure is calculated from five items on the survey. Participants were asked to rate the difficulty of the following: • Effectively communicating with my partner about my intentions and wishes about sexual activity • Refusing unwanted and/or unprotected sex • Identifying if a relationship is healthy or unhealthy • Following the steps for correct condom use • Getting/buying condoms or other birth control The variable is constructed as a continuous variable by taking the average of items answered. Values range from "0" (Very Difficult) to "4" (Very Easy). If respondents answered some items but skipped others, only valid, non-missing responses are counted in the calculations. When respondents skipped all questions in the section, their score is coded as missing. The composite score based on 5 items have an alpha of 0.66 at baseline and 0.66 at 1-year follow-up.	1 year after baseline
Intention to have sex	This variable is a measure of whether a student intends to have sexual intercourse in the next 12 months. The measure is based on the following item in the survey: • Do you intend to have sexual intercourse in the next 12 months? Responses range from "1" (Definite Intention) to "4" (No Intention). A higher score means the student has lower intent to engage in sexual intercourse.	1 year after baseline

Outcome Name	Description of Outcome	Timing of Measure Relative to Program
Intention to use condom during intercourse	This variable is a measure of whether a student intends to use a condom assuming the student has sexual intercourse in the next 12 months. The measure is based on the following item in the survey: • If you were to have sexual intercourse in the next 12 months, do you intend to use (or have your partner use) a condom? Responses are reverse coded, and range from "1" (Definite Intention) to "4" (No Intention). A higher score means the student has higher intent to use a condom during intercourse.	1 year after baseline
Intention to use any birth control method during intercourse	This variable is a measure of whether a student (or his/her partner) intends to use birth control assuming the student has sexual intercourse in the next 12 months. The measure is based on the following item in the survey: • If you were to have sexual intercourse in the next 12 months, do you intend to use (or have your partner use) any of these methods of birth control: birth control pills, the shot (Depo Provera), the patch, the ring (NuvaRing), IUD (Mirena or Paragard), implant (Implanon)? Responses are reverse coded, and range from "1" (Definite Intention) to "4" (No Intention). A higher score means the student has higher intent to use birth control during intercourse.	1 year after baseline

Exhibit C.5.1: School and Student Sample Sizes by Intervention Status

	Time period	Sample Size Total	Sample Size Treatment	Sample Size Control	Response Rate Total	Response Rate Treatment	Response Rate Control
Number of Schools							
0. At beginning of study, originally recruited sample	Random assignment	36	18	18			
1. At beginning of study, study sample excluding withdrawn schools	Random assignment	34	17	17			
2. Contributed at least one student at baseline	Baseline	34	17	17	94.4% of row 0 100.0% of row 1	94.4% of row 0 100.0% of 1	94.4% of row 0 100.0% of 1
3. Contributed at least one student at follow-up	1-year post-baseline	34	17	17	94.4% of row 0 100.0% of row 1	94.4% of row 0 100.0% of 1	94.4% of row 0 100.0% of 1
Number of Students							
4. Study sample (34 schools), excluding withdrawn schools	Random assignment	2,203	1,383	820			
5. Parent consented	See note[(a)]	1,783	1,158	625	80.9% of row 4 100.0% of row 5	83.5% of row 4 100.0% of row 5	76.2% of row 4 100.0% of row 5
6. Contributed a baseline survey (Responded to at least one item on the survey)	Baseline	1,735	1,135	600	78.8% of row 4 97.3% of row 5	82.1% of row 4 98.0% of row 5	73.2% of row 4 96.0% of row 5
7. Contributed a follow-up survey (Responded to at least one item on the survey)	1-yearpost-baseline	1,548	997	551	70.3% of row 4 86.8% of row 5	72.1% of row 4 86.1% of row 5	67.2% of row 4 88.2% of row 5
8. Analytic sample for high risk sexual behaviors (the outcome variable is non-missing)	1-year post-baseline	1,494	961	533	67.8% of row 4 83.8% of row 5	69.5% of row 4 83.0% of row 5	65.0% of row 4 85.3% of row 5
9. Analytic sample for initiation of sex (the outcome variable is non-missing)	1-year post-baseline	1,488	958	530	67.5% of row 4 83.5% of row 5	69.3% of row 4 82.7% of row 5	64.6% of row 4 84.8% of row 5
10. Analytic sample for knowledge of TPP and STI prevention (the outcome variable is non-missing)	1-year post-baseline	1,546	995	551	70.2% of row 4 86.7% of row 5	71.9% of row 4 85.9% of row 5	67.2% of row 4 88.2% of row 5

	Time period	Sample Size Total	Sample Size Treatment	Sample Size Control	Response Rate Total	Response Rate Treatment	Response Rate Control
11. Analytic sample for attitudes toward healthy sexual behaviors (the outcome variable is non-missing)	1-year post-baseline	1,537	990	547	69.8% of row 4 86.2% of row 5	71.6% of row 4 85.5% of row 5	66.7% of row 4 87.5% of row 5
12. Analytic sample for skills to manage relationships and choices (the outcome variable is non-missing)	1-year post-baseline	1,425	943	482	64.7% of row 4 79.9% of row 5	68.2% of row 4 81.4% of row 5	58.8% of row 4 77.1% of row 5
13. Analytic sample for intention to have sex (the outcome variable is non-missing)	1-year post-baseline	1,374	902	472	62.4% of row 4 77.1% of row 5	65.2% of row 4 77.9% of row 5	57.6% of row 4 75.5% of row 5
14. Analytic sample for intention to use condom while having sex (the outcome variable is non-missing)	1-year post-baseline	1,361	911	450	61.8% of row 4 76.3% of row 5	65.9% of row 4 78.7% of row 5	54.9% of row 4 72.0% of row 5
15. Analytic sample for intention to use birth control while having sex (the outcome variable is non-missing)	1-year post-baseline	1,196	823	373	54.3% of row 4 67.1% of row 5	59.5% of row 4 71.1% of row 5	45.5% of row 4 59.7% of row 5

Note: The study was originally designed to collect the 2-year follow-up. Due to unforeseen circumstances, the second year data collection was interrupted, and only partial data could be collected under conditions significantly differently from original data collection protocols. Overall, of 1,771 students with parental consent for second-year follow-up, 579 completed the survey. Due to incomplete data collection, this report does not contain analyses of second-year data.

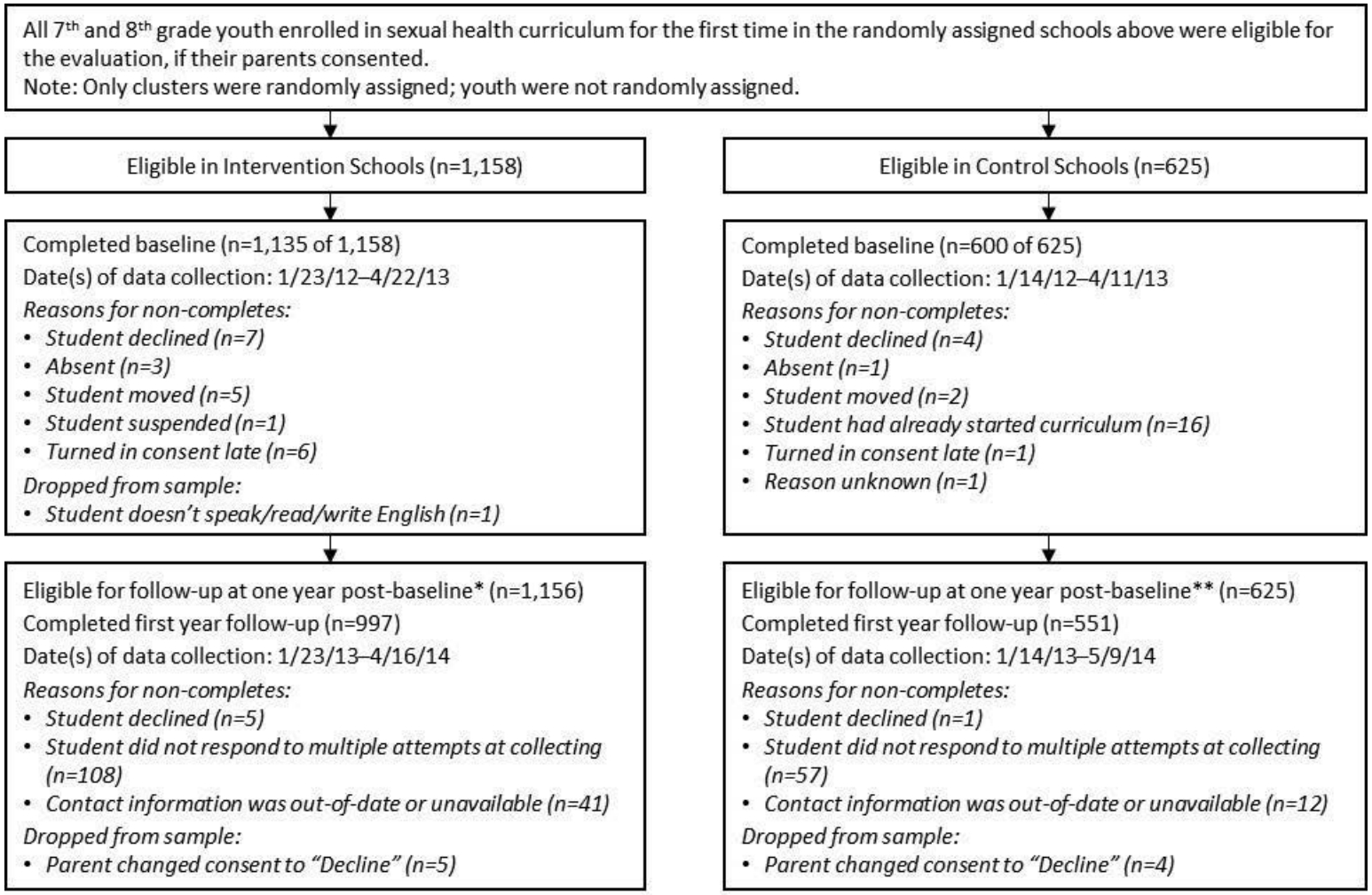
(a) Parental consent was obtained prior to baseline data collection. Parents were kept from knowing the assignment condition before they returned consent. Students were excluded from any data collection activity if parents did not consent to their participation in the study. Assent from students with parental consent was obtained at the time of each data collection (students were allowed to not take a survey even if parents consented). Non-assent by students is counted as non-response.

Exhibit C.5.2: School and Student Sample Sizes by Intervention Status, for the Alternative Sample (Non-Attriting Blocks Only)

	Sample Size Time period	Sample Size Total	Sample Size Treatment	Sample Size Control	Response Rate Total	Response Rate Treatment	Response Rate Control
Number of Schools							
1. At beginning of study, alternative sample, excluding blocks that included withdrawn schools	Random assignment	31	15	16			
2. Contributed at least one student at baseline	Baseline	31	15	16	100.0% of row 1	100.0% of row 1	100.0% of row 1
3. Contributed at least one student at follow-up	12 months after baseline	31	15	16	100.0% of row 1	100.0% of row 1	100.0% of row 1
Number of Students							
4. In non-attriting schools, in alternative sample, excluding the blocks with withdrawn schools	Random assignment	2,047	1,195	652			
5. Parent consented	See note[(a)]	1,656	1,083	573			
6. Contributed a baseline survey (Responded to at least one item on the survey. Not used for impact analyses)	Baseline	1,622	1,070	552	79.2% of row 4 97.9% of row 5	89.5% of row 4 98.8% of row 5	84.7% of row 4 96.3% of row 4
7. Contributed a follow-up survey (Responded to at least one item on the survey)	1-year post-baseline	1,459	946	513	71.3% of row 4 88.1% of row 5	79.2% of row 4 87.3% of row 5	78.7% of row 4 89.5% of row 5

(a) Parental consent was obtained prior to baseline data collection. Parents were kept from knowing the assignment condition before they returned consent. Students were excluded from any data collection activity if parents did not consent to their participation in the study. Assent from students with parental consent was obtained at the time of each data collection (students were allowed to not to take a survey even if parents consented). Non-assent by students is counted as non-response.

Exhibit C.6: CONSORT Diagram for Pono Choices Student Sample



*Eligible students include all students whose parents consented to the study, except for one student whose parents later changed their consent and one student who did not speak, read or write English.
** Eligible students include all students whose parents consented to the study.

Exhibit C.7: Pre-Treatment Characteristics of Students at Baseline: Total Baseline Sample

Variable	All* Obs[a]	All Mean	All Std. Dev.	Treat-ment** Obs[a]	Treat-ment Mean	Treat-ment Std. Dev.	Control *** Obs[a]	Control Mean	Control Std. Dev.
Outcome measures at baseline[b]									
Engagement in high-risk sexual behavior ("1"=used neither condom or birth control when having sex, "0"=otherwise)	1,529	0.01	0.08	1,005	0.01	0.08	524	0.00	0.06
Initiation of sexual activity (1=ever had sex, 0=otherwise)	1,489	0.07	0.26	981	0.07	0.25	508	0.09	0.29
Knowledge of pregnancy and STI prevention (proportion of correct responses to 10 questions)	1,726	0.62	0.20	1,129	0.63	0.19	597	0.60	0.21
Attitudes toward healthy sexual behaviors (average score on scale of 1-4 where 4=very important)	1,724	3.40	0.54	1,126	3.42	0.50	598	3.37	0.60
Skills in managing relationships and choices (average score on scale of 1-4 where 1=very difficult and 4=very easy)	1,689	2.78	0.60	1,109	2.78	0.58	580	2.79	0.62
Intention to have sex (average score on scale of 1-4 where 4=very likely)	1,651	3.35	0.84	1,085	3.34	0.84	566	3.35	0.85
Intention to use condom during intercourse (average score on scale of 1-4 where 4=very likely)	1,626	3.49	0.88	1,066	3.51	0.85	560	3.44	0.95
Intention to use birth control during intercourse (average score on scale of 1-4 where 4=very likely)	1,583	3.25	0.92	1,038	3.25	0.90	545	3.24	0.96
Student characteristics at baseline									
Age (years)	1,734	12.31	0.64	1,129	12.30	0.65	605	12.32	0.64
Female ("1"=female, "0"=Otherwise)	1,742	0.52	0.50	1,133	0.51	0.50	609	0.54	0.50
Bisexual/homosexual orientation ("1"=Bi/homosexual; "0"=Otherwise)	1,724	0.03	0.16	1,123	0.03	0.17	601	0.02	0.15
Parent does not speak English ("1"=Does not speak English; "0"=Otherwise)	1,752	0.02	0.13	1,139	0.02	0.13	613	0.02	0.14
Student using non-English at home ("1"=Use non English at home; "0"=Otherwise)[e]	1,767	0.31	0.46	1,149	0.27	0.45	618	0.39	0.49
Receiving mostly As and Bs for grade ("1"=As and Bs, "0"=Otherwise)	1,654	0.71	0.46	1,078	0.71	0.46	576	0.71	0.46

Variable	All* Obs[a]	All Mean	All Std. Dev.	Treat-ment** Obs[a]	Treat-ment Mean	Treat-ment Std. Dev.	Control *** Obs[a]	Control Mean	Control Std. Dev.
Grade level at start of study ("1"=7th grade, "0"=otherwise/8th grade)	1,755	0.87	0.34	1,140	0.89	0.32	615	0.84	0.37
Native Hawaiian ("1"=Native Hawaiian including mixed race, "0"=Otherwise)[c]	1,745	0.43	0.50	1,131	0.41	0.49	614	0.49	0.50
Asian ("1"=Asian including mixed race, "0"=Otherwise)[c]	1,745	0.71	0.45	1,131	0.70	0.46	614	0.73	0.45
African American ("1"=African American including mixed race, "0"=otherwise)[c]	1,745	0.06	0.24	1,131	0.07	0.25	614	0.06	0.23
Hispanic ("1"=Hispanic, "0"=Otherwise) [d]	1,635	0.18	0.38	1,062	0.18	0.38	573	0.18	0.39
Mutually exclusive race/ethnicity classifications:									
Non-Hispanic Hawaiian ("1"=Native Hawaiian including no mixed race, "0"=Otherwise)	1,624	0.02	0.15	1,053	0.02	0.14	571	0.03	0.17
Non-Hispanic Asian ("1"Asian including no mixed race, "0"=Otherwise)	1,624	0.20	0.40	1,053	0.19	0.39	571	0.21	0.41
Non-Hispanic African American ("1"=African American including mixed race, "0"=Otherwise)	1,624	0.01	0.09	1,053	0.01	0.10	571	0.01	0.07
Non-Hispanic White ("1"=White including mixed race, "0"=Otherwise)	1,624	0.08	0.26	1,053	0.08	0.27	571	0.07	0.25

Source: Student Baseline Survey

* All N=1767

** Treatment N=1,146

*** Control N=621

(a) The number of students responding to the relevant question on the baseline survey. Demographic information was collected from 1,767 students (1,735 of these students completed a baseline survey while an additional 32 students did not complete a baseline survey but provided demographic information during follow-up).

(b) See Table 4 in the main text and Exhibit C.4 in Appendix C for the definition of outcome measures.

(c) The number of participants who selected at least one option for race and exclude those who did not select any race categories. The total percentages exceed 100% as many students reported two or more races.

(d) Only students for whom we know their definite Hispanic status are included when calculating percentages. Students that have unknown Hispanic status are excluded from the percentage.

(e) The number of participants who selected at least one option for language spoken at home. The numbers exclude students who did not select any language option, and are assumed to have skipped the question completely.

Exhibit C.8: Summary Statistics of Key Baseline Measures of Student Characteristics
Analytic Sample for Estimating Impact on Engagement in High-Risk Sexual Behavior

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Engagement in high-risk sexual behavior ("1"=used neither condom or birth control when having sex, "0"=otherwise)	843	0.01	0.08	453	0.00	0.05	0.00	0.160	X
Age in years	943	12.27	0.64	521	12.28	0.60	-0.01	0.589	X
Grade level at start of study ("1"=7th grade, "0"=otherwise/8th grade)	953	0.90	0.30	530	0.86	0.35	0.04	0.426	
Female ("1"=female, "0"=Otherwise)	949	0.52	0.50	525	0.54	0.50	-0.02	0.559	X
Bisexual/homosexual ("1"=Bi/homosexual; "0"=Otherwise)	937	0.03	0.16	519	0.02	0.14	0.01	0.672	
Student using non-English at home ("1"=Use non English at home; "0"=Otherwise)	961	0.27	0.44	533	0.39	0.49	-0.12**	0.000	X
Parent does not speak English ("1"=Does not speak English; "0"=Otherwise)	951	0.02	0.12	529	0.02	0.15	-0.01	0.051	
Receiving mostly As and Bs for grade ("1"=As and Bs, "0"=Otherwise)	904	0.72	0.45	499	0.71	0.45	0.01	0.726	X
Native Hawaiian ("1"=Native Hawaiian including mixed race, "0"=Otherwise)	948	0.41	0.49	529	0.49	0.50	-0.08	0.099	X
Asian ("1"=Asian including mixed race, "0"=Otherwise)	948	0.72	0.45	529	0.74	0.44	-0.02*	0.032	X

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Mutually exclusive race/ethnicity classifications:									
Non-Hispanic Hawaiian ("1"=Native Hawaiian including no mixed race, "0"=Otherwise)	886	0.02	0.13	492	0.03	0.17	-0.01	0.246	
Non-Hispanic Asian ("1"Asian including no mixed race, "0"=Otherwise)	886	0.20	0.40	492	0.22	0.42	-0.02	0.438	
Non-Hispanic African American ("1"= African American including no mixed race, "0"=Otherwise)	886	0.01	0.10	492	0.01	0.08	0.00	0.726	
Non-Hispanic White ("1"=White including mixed race, "0"=Otherwise)	886	0.08	0.27	492	0.06	0.24	0.01	0.135	
Hispanic ("1"=Hispanic, "0"=Otherwise	890	0.18	0.38	494	0.17	0.38	0.01	0.239	

Source: Authors' calculations based on student surveys.

**Significant at the .01 level. * Significant at the .05 level.

The group difference was evaluated applying the same statistical model used to estimate the impact, with only the assignment group and blocking variables as controls. The size of the analytic sample is 1494 (961 in intervention and 533 in control). Equivalence tests were conducted using non-missing observations of baseline variables. In the impact estimation, the dummy variable adjustment method was used for missing baseline variables as covariates.

For race/ethnic indicators used as covariates in the study, we counted a student in a particular race/ethnic group if s/he identified himself/herself as that racial/ethnic group, regardless of whether the student identified with one or more other race/ethnic groups. In addition, we also examined the racial/ethnic distribution based on the conventional U.S. Census definition of race/ethnic groups, where a group is defined excluding mixed race individuals. These mutual exclusive race/ethic categories are presented in the bottom rows of the table.

Exhibit C.9: Summary Statistics of Key Baseline Measures for Student Characteristics
Analytic Sample for Estimating Impact on Initiation of Sexual Activity

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Initiation of sexual activity (1=ever had sex, 0=otherwise)	825	0.06	0.25	440	0.07	0.26	-0.01	0.205	X
Age in years	941	12.27	0.64	518	12.28	0.60	-0.01	0.599	X
Grade level at start of study ("1"=7th grade, "0"=otherwise/8th grade)	950	0.90	0.30	527	0.86	0.35	0.04	0.423	
Female ("1"=female, "0"=Otherwise)	946	0.53	0.50	522	0.55	0.50	-0.02	0.485	X
Bisexual/homosexual ("1"=Bi/homosexual; "0"=Otherwise)	936	0.03	0.16	516	0.02	0.14	0.01	0.686	
Student using non-English at home ("1"=Use non English at home; "0"=Otherwise)	958	0.27	0.44	530	0.39	0.49	-0.12**	0.000	X
Parent does not speak English ("1"=Does not speak English; "0"=Otherwise)	948	0.02	0.12	526	0.02	0.14	-0.01	0.078	
Receiving mostly As and Bs for grade ("1"=As and Bs, "0"=Otherwise)	900	0.72	0.45	496	0.71	0.45	0.01	0.719	X
Native Hawaiian ("1"=Native Hawaiian including mixed race, "0"=Otherwise)	945	0.41	0.49	526	0.49	0.50	-0.08	0.109	X
Asian ("1"=Asian including mixed race, "0"=Otherwise)	945	0.71	0.45	526	0.74	0.44	-0.03*	0.024	X

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Mutually exclusive race/ ethnicity classifications:									
Non-Hispanic Hawaiian ("1"=Native Hawaiian including no mixed race, "0"=Otherwise)	883	0.02	0.13	490	0.03	0.17	-0.01	0.241	
Non-Hispanic Asian ("1"Asian including no mixed race, "0"=Otherwise)	883	0.20	0.40	490	0.22	0.42	-0.03	0.323	
Non-Hispanic African American ("1"= African American including no mixed race, "0"=Otherwise)	883	0.01	0.10	490	0.01	0.08	0.00	0.725	
Non-Hispanic White ("1"=White including mixed race, "0"=Otherwise)	883	0.08	0.27	490	0.06	0.24	0.01	0.130	
Hispanic ("1"=Hispanic, "0"=Otherwise	887	0.18	0.38	492	0.17	0.38	0.01	0.251	

Source: Authors' calculations based on student surveys.

**Significant at the .01 level. * Significant at the .05 level.

The group difference was evaluated applying the same statistical model used to estimate the impact, with only the assignment group and blocking variables as controls. The size of the analytic sample is 1488 (958 in intervention and 530 in control). Equivalence tests were conducted using non-missing observations of baseline variables. In the impact estimation, the dummy variable adjustment method was used for missing baseline variables as covariates.

For race/ethnic indicators used as covariates in the study, we counted a student in a particular race/ethnic group if s/he identified himself/herself as that racial/ethnic group, regardless of whether the student identified with one or more other race/ethnic groups. In addition, we also examined the racial/ethnic distribution based on the conventional U.S. Census definition of race/ethnic groups, where a group is defined excluding mixed race individuals. These mutual exclusive race/ethic categories are presented in the bottom rows of the table.

Exhibit C.10: Summary Statistics of Key Baseline Measures for Student Characteristics
Analytic Sample for Estimating Impact on Knowledge about Pregnancy and STI Prevention

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Knowledge of pregnancy and STI prevention (proportion of correct responses to 10 questions	977	0.62	0.19	530	0.60	0.21	0.02	0.166	X
Age in years	977	12.27	0.64	539	12.28	0.60	-0.01	0.634	X
Grade level at start of study ("1"=7th grade, "0"=otherwise/8th grade)	988	0.90	0.30	548	0.86	0.35	0.04	0.458	
Female ("1"=female, "0"=Otherwise)	984	0.52	0.50	543	0.55	0.50	-0.02	0.606	X
Bisexual/homosexual ("1"=Bi/homosexual; "0"=Otherwise)	972	0.03	0.16	536	0.02	0.15	0.00	0.879	
Student using non-English at home ("1"=Use non English at home; "0"=Otherwise)	995	0.27	0.44	551	0.39	0.49	-0.13**	0.000	X
Parent does not speak English ("1"=Does not speak English; "0"=Otherwise)	986	0.02	0.13	547	0.02	0.15	-0.01	0.071	
Receiving mostly As and Bs for grade ("1"=As and Bs, "0"=Otherwise)	937	0.73	0.45	514	0.71	0.46	0.02	0.613	X
Native Hawaiian ("1"=Native Hawaiian including mixed race, "0"=Otherwise)	982	0.41	0.49	547	0.49	0.50	-0.08	0.084	X
Asian ("1"=Asian including mixed race, "0"=Otherwise)	982	0.72	0.45	547	0.74	0.44	-0.03*	0.023	X

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Mutually exclusive race/ethnicity classifications:									
Non-Hispanic Hawaiian ("1"=Native Hawaiian including no mixed race, "0"=Otherwise)	917	0.02	0.13	509	0.03	0.16	-0.01	0.256	
Non-Hispanic Asian ("1"Asian including no mixed race, "0"=Otherwise)	917	0.20	0.40	509	0.22	0.42	-0.03	0.309	
Non-Hispanic African American ("1"= African American including no mixed race, "0"=Otherwise)	917	0.01	0.10	509	0.01	0.08	0.00	0.729	
Non-Hispanic White ("1"=White including mixed race, "0"=Otherwise)	917	0.08	0.27	509	0.06	0.24	0.02	0.112	
Hispanic ("1"=Hispanic, "0"=Otherwise	921	0.17	0.38	511	0.17	0.38	0.00	0.229	

Source: Authors' calculations based on student surveys.

**Significant at the .01 level. * Significant at the .05 level.

The group difference was evaluated applying the same statistical model used to estimate the impact, with only the assignment group and blocking variables as controls. The size of the analytic sample is 1546 (995 in intervention and 551 in control). Equivalence tests were conducted using non-missing observations of baseline variables. In the impact estimation, the dummy variable adjustment method was used for missing baseline variables as covariates.

For race/ethnic indicators used as covariates in the study, we counted a student in a particular race/ethnic group if s/he identified himself/herself as that racial/ethnic group, regardless of whether the student identified with one or more other race/ethnic groups. In addition, we also examined the racial/ethnic distribution based on the conventional U.S. Census definition of race/ethnic groups, where a group is defined excluding mixed race individuals. These mutual exclusive race/ethic categories are presented in the bottom rows of the table.]

Exhibit C.11: Summary Statistics of Key Baseline Measures for Student Characteristics
Analytic Sample for Estimating Impact on Attitudes toward Healthy Sexual Behavior

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Attitudes toward healthy sexual behaviors (average score on scale of 1-4 where 4=very important)	972	3.428	0.507	529	3.367	0.610	0.061*	0.023	X
Age in years	972	12.27	0.64	535	12.28	0.60	-0.01	0.642	X
Grade level at start of study ("1"=7th grade, "0"=otherwise/8th grade)	983	0.90	0.30	544	0.86	0.35	0.04	0.429	
Female ("1"=female, "0"=Otherwise)	979	0.52	0.50	539	0.55	0.50	-0.02	0.621	X
Bisexual/homosexual ("1"=Bi/homosexual; "0"=Otherwise)	967	0.02	0.16	532	0.02	0.15	0.00	0.974	
Student using non-English at home ("1"=Use non English at home; "0"=Otherwise)	990	0.27	0.44	547	0.39	0.49	-0.12**	0.000	X
Parent does not speak English ("1"=Does not speak English; "0"=Otherwise)	981	0.02	0.12	543	0.02	0.15	-0.01*	0.039	
Receiving mostly As and Bs for grade ("1"=As and Bs, "0"=Otherwise)	932	0.73	0.45	511	0.71	0.45	0.02	0.607	X
Native Hawaiian ("1"=Native Hawaiian including mixed race, "0"=Otherwise)	977	0.41	0.49	543	0.49	0.50	-0.08	0.086	X
Asian ("1"=Asian including mixed race, "0"=Otherwise)	977	0.72	0.45	543	0.74	0.44	-0.02*	0.025	X

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Mutually exclusive race/ ethnicity classifications:									
Non-Hispanic Hawaiian ("1"=Native Hawaiian including no mixed race, "0"=Otherwise)	912	0.02	0.13	505	0.03	0.16	-0.01	0.256	
Non-Hispanic Asian ("1"Asian including no mixed race, "0"=Otherwise)	912	0.20	0.40	505	0.22	0.42	-0.03	0.314	
Non-Hispanic African American ("1"= African American including no mixed race, "0"=Otherwise)	912	0.01	0.10	505	0.01	0.08	0.00	0.729	
Non-Hispanic White ("1"=White including mixed race, "0"=Otherwise)	912	0.08	0.27	505	0.06	0.24	0.02	0.106	
Hispanic ("1"=Hispanic, "0"=Otherwise	916	0.18	0.38	507	0.17	0.38	0.00	0.200	

Source: Authors' calculations based on student surveys.

**Significant at the .01 level. * Significant at the .05 level.

The group difference was evaluated applying the same statistical model used to estimate the impact, with only the assignment group and blocking variables as controls. The size of the analytic sample is 1537 (990 in intervention and 547 in control). Equivalence tests were conducted using non-missing observations of baseline variables. In the impact estimation, the dummy variable adjustment method was used for missing baseline variables as covariates.

For race/ethnic indicators used as covariates in the study, we counted a student in a particular race/ethnic group if s/he identified himself/herself as that racial/ethnic group, regardless of whether the student identified with one or more other race/ethnic groups. In addition, we also examined the racial/ethnic distribution based on the conventional U.S. Census definition of race/ethnic groups, where a group is defined excluding mixed race individuals. These mutual exclusive race/ethic categories are presented in the bottom rows of the table.

Exhibit C.12: Summary Statistics of Key Baseline Measures for Student Characteristics
Analytic Sample for Estimating Impact on Skills in Managing Relationships and Choices

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Skills in managing relationships and choices (average score on scale of 1-4 where 1=very difficult and 4=very easy)	912	2.78	0.57	456	2.78	0.60	0.00	0.971	X
Age in years	912	2.78	0.57	456	2.78	0.60	0.00	0.971	X
Grade level at start of study ("1"=7th grade, "0"=otherwise/8th grade)	925	12.27	0.64	470	12.29	0.60	-0.02	0.582	
Female ("1"=female, "0"=Otherwise)	935	0.90	0.30	479	0.85	0.36	0.05	0.387	X
Bisexual/homosexual ("1"=Bi/homosexual; "0"=Otherwise)	931	0.52	0.50	476	0.54	0.50	-0.01	0.854	
Student using non-English at home ("1"=Use non English at home; "0"=Otherwise)	921	0.03	0.16	470	0.03	0.16	0.00	0.831	X
Parent does not speak English ("1"=Does not speak English; "0"=Otherwise)	943	0.26	0.44	482	0.39	0.49	-0.13**	0.000	
Receiving mostly As and Bs for grade ("1"=As and Bs, "0"=Otherwise)	934	0.02	0.13	478	0.03	0.16	-0.01*	0.024	X
Native Hawaiian ("1"=Native Hawaiian including mixed race, "0"=Otherwise)	887	0.73	0.44	449	0.70	0.46	0.03	0.474	X

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Asian ("1"=Asian including mixed race, "0"=Otherwise)	931	0.41	0.49	478	0.50	0.50	-0.09	0.056	X
Mutually exclusive race/ ethnicity classifications:									
Non-Hispanic Hawaiian ("1"=Native Hawaiian including no mixed race, "0"=Otherwise)	873	0.02	0.13	449	0.03	0.16	-0.01	0.370	
Non-Hispanic Asian ("1"Asian including no mixed race, "0"=Otherwise)	873	0.20	0.40	449	0.22	0.41	-0.02	0.313	
Non-Hispanic African American ("1"= African American including no mixed race, "0"=Otherwise)	873	0.01	0.10	449	0.01	0.08	0.00	0.829	
Non-Hispanic White ("1"=White including mixed race, "0"=Otherwise)	873	0.08	0.27	449	0.06	0.24	0.02	0.119	
Hispanic ("1"=Hispanic, "0"=Otherwise	877	0.18	0.38	451	0.18	0.38	0.00	0.279	

Source: Authors' calculations based on student surveys.

**Significant at the .01 level. * Significant at the .05 level.

The group difference was evaluated applying the same statistical model used to estimate the impact, with only the assignment group and blocking variables as controls. The size of the analytic sample is 1425 (943 in intervention and 482 in control). Equivalence tests were conducted using non-missing observations of baseline variables. In the impact estimation, the dummy variable adjustment method was used for missing baseline variables as covariates.

For race/ethnic indicators used as covariates in the study, we counted a student in a particular race/ethnic group if s/he identified himself/herself as that racial/ethnic group, regardless of whether the student identified with one or more other race/ethnic groups. In addition, we also examined the racial/ethnic distribution based on the conventional U.S. Census definition of race/ethnic groups, where a group is defined excluding mixed race individuals. These mutual exclusive race/ethic categories are presented in the bottom rows of the table.

Exhibit C.13: Summary Statistics of Key Baseline Measures for Student Characteristics
Analytic Sample for Estimating Impact on Intent to Have Sex

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Intention to have sex (average score on scale of 1-4 where 4=very likely)	856	3.39	0.80	437	3.41	0.82	-0.01	0.922	X
Age in years	885	12.26	0.63	463	12.28	0.60	-0.02	0.655	X
Grade level at start of study ("1"=7th grade, "0"=otherwise/8th grade)	894	0.91	0.29	469	0.86	0.35	0.06	0.456	
Female ("1"=female, "0"=Otherwise)	890	0.54	0.50	467	0.57	0.50	-0.03	0.309	X
Bisexual/homosexual ("1"=Bi/homosexual; "0"=Otherwise)	882	0.03	0.16	459	0.02	0.15	0.00	0.675	
Student using non-English at home ("1"=Use non English at home; "0"=Otherwise)	902	0.26	0.44	472	0.38	0.48	-0.11**	0.000	X
Parent does not speak English ("1"=Does not speak English; "0"=Otherwise)	894	0.02	0.12	469	0.02	0.14	0.00	0.181	
Receiving mostly As and Bs for grade ("1"=As and Bs, "0"=Otherwise)	850	0.73	0.44	440	0.73	0.45	0.01	0.704	X
Native Hawaiian ("1"=Native Hawaiian including mixed race, "0"=Otherwise)	891	0.41	0.49	468	0.50	0.50	-0.09	0.077	X
Asian ("1"=Asian including mixed race, "0"=Otherwise)	891	0.71	0.45	468	0.75	0.43	-0.04*	0.011	X

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Mutually exclusive race/ ethnicity classifications:									
Non-Hispanic Hawaiian ("1"=Native Hawaiian including no mixed race, "0"=Otherwise)	838	0.02	0.14	437	0.03	0.16	0.00	0.660	
Non-Hispanic Asian ("1"Asian including no mixed race, "0"=Otherwise)	838	0.20	0.40	437	0.23	0.42	-0.04	0.191	
Non-Hispanic African American ("1"= African American including no mixed race, "0"=Otherwise)	838	0.01	0.09	437	0.01	0.08	0.00	0.882	
Non-Hispanic White ("1"=White including mixed race, "0"=Otherwise)	838	0.08	0.26	437	0.06	0.23	0.02	0.073	
Hispanic ("1"=Hispanic, "0"=Otherwise	842	0.18	0.38	439	0.18	0.38	0.00	0.275	

Source: Authors' calculations based on student surveys.

**Significant at the .01 level. * Significant at the .05 level.

The group difference was evaluated applying the same statistical model used to estimate the impact, with only the assignment group and blocking variables as controls. The size of the analytic sample is 1374 (902 in intervention and 472 in control). Equivalence tests were conducted using non-missing observations of baseline variables. In the impact estimation, the dummy variable adjustment method was used for missing baseline variables as covariates.

For race/ethnic indicators used as covariates in the study, we counted a student in a particular race/ethnic group if s/he identified himself/herself as that racial/ethnic group, regardless of whether the student identified with one or more other race/ethnic groups. In addition, we also examined the racial/ethnic distribution based on the conventional U.S. Census definition of race/ethnic groups, where a group is defined excluding mixed race individuals. These mutual exclusive race/ethic categories are presented in the bottom rows of the table.

Exhibit C.14: Summary Statistics of Key Baseline Measures for Student Characteristics
Analytic Sample for Estimating Impact on Intent to Use a Condom during Intercourse

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Intention to use condom during intercourse (average score on scale of 1-4 where 4=very likely)	854	3.57	0.79	416	3.52	0.88	0.05	0.045	X
Age in years	894	12.27	0.63	442	12.28	0.59	-0.01	0.740	X
Grade level at start of study ("1"=7th grade, "0"=otherwise/8th grade)	904	0.91	0.29	447	0.84	0.36	0.06	0.450	
Female ("1"=female, "0"=Otherwise)	901	0.52	0.50	444	0.55	0.50	-0.03	0.514	X
Bisexual/homosexual ("1"=Bi/homosexual; "0"=Otherwise)	889	0.03	0.16	439	0.03	0.16	0.00	0.860	
Student using non-English at home ("1"=Use non English at home; "0"=Otherwise)	911	0.26	0.44	450	0.38	0.48	-0.11**	0.000	X
Parent does not speak English ("1"=Does not speak English; "0"=Otherwise)	902	0.01	0.12	447	0.02	0.14	-0.01	0.109	
Receiving mostly As and Bs for grade ("1"=As and Bs, "0"=Otherwise)	859	0.73	0.44	419	0.70	0.46	0.03	0.455	X
Native Hawaiian ("1"=Native Hawaiian including mixed race, "0"=Otherwise)	899	0.41	0.49	446	0.51	0.50	-0.11	0.058	X

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Asian ("1"=Asian including mixed race, "0"=Otherwise)	899	0.72	0.45	446	0.75	0.43	-0.03*	0.020	X
Mutually exclusive race/ ethnicity classifications:									
Non-Hispanic Hawaiian ("1"=Native Hawaiian including no mixed race, "0"=Otherwise)	842	0.02	0.13	418	0.03	0.17	-0.01	0.350	
Non-Hispanic Asian ("1"Asian including no mixed race, "0"=Otherwise)	842	0.19	0.40	418	0.21	0.41	-0.02	0.381	
Non-Hispanic African American ("1"= African American including no mixed race, "0"=Otherwise)	842	0.01	0.10	418	0.01	0.08	0.00	0.841	
Non-Hispanic White ("1"=White including mixed race, "0"=Otherwise)	842	0.08	0.28	418	0.06	0.25	0.02	0.095	
Hispanic ("1"=Hispanic, "0"=Otherwise	846	0.17	0.38	420	0.18	0.38	0.00	0.389	

Source: Authors' calculations based on student surveys.

**Significant at the .01 level. * Significant at the .05 level.

The group difference was evaluated applying the same statistical model used to estimate the impact, with only the assignment group and blocking variables as controls. The size of the analytic sample is 1361 (911 in intervention and 450 in control). Equivalence tests were conducted using non-missing observations of baseline variables. In the impact estimation, the dummy variable adjustment method was used for missing baseline variables as covariates.

For race/ethnic indicators used as covariates in the study, we counted a student in a particular race/ethnic group if s/he identified himself/herself as that racial/ethnic group, regardless of whether the student identified with one or more other race/ethnic groups. In addition, we also examined the racial/ethnic distribution based on the conventional U.S. Census definition of race/ethnic groups, where a group is defined excluding mixed race individuals. These mutual exclusive race/ethic categories are presented in the bottom rows of the table.

Exhibit C.15: Summary Statistics of Key Baseline Measures for Student Characteristics
Analytic Sample for Estimating Impact on Intent to Use Birth Control during Intercourse

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Intention to use birth control during intercourse (average score on scale of 1-4 where 4=very likely)	754	3.35	0.84	337	3.35	0.89	-0.01	0.372	X
Age in years	807	12.28	0.62	368	12.28	0.58	-0.01	0.795	X
Grade level at start of study ("1"=7th grade, "0"=otherwise/8th grade)	816	0.91	0.29	370	0.84	0.37	0.07	0.367	
Female ("1"=female, "0"=Otherwise)	814	0.53	0.50	369	0.55	0.50	-0.02	0.692	X
Bisexual/homosexual ("1"=Bi/homosexual; "0"=Otherwise)	803	0.03	0.16	363	0.03	0.16	0.00	0.664	
Student using non-English at home ("1"=Use non English at home; "0"=Otherwise)	823	0.26	0.44	373	0.36	0.48	-0.11**	0.000	X
Parent does not speak English ("1"=Does not speak English; "0"=Otherwise)	815	0.01	0.11	371	0.02	0.14	-0.01	0.081	
Receiving mostly As and Bs for grade ("1"=As and Bs, "0"=Otherwise)	771	0.75	0.43	344	0.72	0.45	0.03	0.528	X
Native Hawaiian ("1"=Native Hawaiian including mixed race, "0"=Otherwise)	811	0.41	0.49	370	0.51	0.50	-0.11	0.053	X

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Asian ("1"=Asian including mixed race, "0"=Otherwise)	811	0.71	0.45	370	0.74	0.44	-0.03*	0.037	X
Mutually exclusive race/ ethnicity classifications:									
Non-Hispanic Hawaiian ("1"=Native Hawaiian including no mixed race, "0"=Otherwise)	762	0.02	0.14	345	0.03	0.18	-0.01	0.225	
Non-Hispanic Asian ("1"Asian including no mixed race, "0"=Otherwise)	762	0.19	0.39	345	0.22	0.42	-0.03	0.277	
Non-Hispanic African American ("1"= African American including no mixed race, "0"=Otherwise)	762	0.01	0.10	345	0.01	0.09	0.00	0.882	
Non-Hispanic White ("1"=White including mixed race, "0"=Otherwise)	762	0.09	0.28	345	0.07	0.25	0.02	0.190	
Hispanic ("1"=Hispanic, "0"=Otherwise	766	0.17	0.38	346	0.18	0.39	-0.01	0.471	

Source: Authors' calculations based on student surveys.

**Significant at the .01 level. * Significant at the .05 level.

The group difference was evaluated applying the same statistical model used to estimate the impact, with only the assignment group and blocking variables as controls. The size of the analytic sample is 1196 (823 in intervention and 373 in control). Equivalence tests were conducted using non-missing observations of baseline variables. In the impact estimation, the dummy variable adjustment method was used for missing baseline variables as covariates.

For race/ethnic indicators used as covariates in the study, we counted a student in a particular race/ethnic group if s/he identified himself/herself as that racial/ethnic group, regardless of whether the student identified with one or more other race/ethnic groups. In addition, we also examined the racial/ethnic distribution based on the conventional U.S. Census definition of race/ethnic groups, where a group is defined excluding mixed race individuals. These mutual exclusive race/ethic categories are presented in the bottom rows of the table.

Exhibit C.16: Summary Statistics of Key Baseline Measures for School Characteristics in Analytic Samples
All Outcome Measures

Measures	Total* N	Total Mean	Total SD	Treat-ment** N	Treatment Mean	Treat-ment SD	Control*** N	Control Mean	Control SD	Difference in Mean	*p*-value for Difference
Classroom size (number of students)	33	14.43	2.97	17	15.29	3.04	16	13.51	2.70	1.79	0.084
School size (total enrollment)	33	528	368	17	611	360	16	440	367	171	0.186
Poverty level (ratio of students eligible for free or reduced-price lunch)	33	0.56	0.18	17	0.55	0.18	16	0.57	0.19	-0.02	0.800
English proficiency (percentage of students classified as English language learners)	31	0.07	0.07	16	0.06	0.05	15	0.07	0.09	-0.01	0.753
Performance (1= not meeting AYP, in restructuring, or a Race to the Top priority school, 0=otherwise)	32	0.06	0.25	17	0.12	0.33	15	0.00	0.00	0.12	0.171

* Total N=34

** Treatment N=17

*** Control N=17

Source: Hawai'i Department of Education and Common Core Data (2010-11 and 2011-12), National Center for Education Statistics.

Exhibit C.17: Summary Statistics of Key Baseline Measures for Student Characteristics Analytic Sample for Estimating Impact on Engagement in High-Risk Sexual Behavior, Excluding Two Blocks that Included Withdrawn Schools (31 schools)

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Engagement in high-risk sexual behavior ("1"=used neither condom or birth control when having sex, "0"=otherwise)	796	0.01	0.08	417	0.00	0.05	0.00	0.253	X
Age in years	895	12.23	0.62	485	12.24	0.58	-0.01	0.635	X
Grade level at start of study ("1"=7th grade, "0"=otherwise/8th grade)	904	0.95	0.22	494	0.89	0.32	0.06	0.426	
Female ("1"=female, "0"=Otherwise)	900	0.52	0.50	489	0.54	0.50	-0.02	0.813	X
Bisexual/homosexual ("1"=Bi/homosexual; "0"=Otherwise)	888	0.02	0.16	483	0.02	0.14	0.01	0.596	
Student using non-English at home ("1"=Use non English at home; "0"=Otherwise)	911	0.26	0.44	497	0.38	0.49	-0.12	0.000	X
Parent does not speak English ("1"=Does not speak English; "0"=Otherwise)	902	0.02	0.12	493	0.02	0.15	-0.01	0.035	
Receiving mostly As and Bs for grade ("1"=As and Bs, "0"=Otherwise)	858	0.72	0.45	464	0.71	0.45	0.01	0.824	X
Native Hawaiian ("1"=Native Hawaiian including mixed race, "0"=Otherwise)	899	0.43	0.50	493	0.49	0.50	-0.06	0.391	X

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Asian ("1"=Asian including mixed race, "0"=Otherwise)	899	0.73	0.44	493	0.75	0.43	-0.02	0.085	X
Mutually exclusive race/ ethnicity classifications:									
Non-Hispanic Hawaiian ("1"=Native Hawaiian including no mixed race, "0"=Otherwise)	837	0.02	0.14	457	0.03	0.17	-0.01	0.246	
Non-Hispanic Asian ("1"Asian including no mixed race, "0"=Otherwise)	837	0.20	0.40	457	0.23	0.42	-0.03	0.268	
Non-Hispanic Black (Includes no other race)	837	0.00	0.06	457	0.00	0.07	0.00	0.913	
Non-Hispanic White ("1"=White including mixed race, "0"=Otherwise)	837	0.07	0.25	457	0.06	0.24	0.01	0.294	
Non-Hispanic with two or more races	837	0.49	0.50	457	0.46	0.50	0.04	0.396	
Hispanic ("1"=Hispanic, "0"=Otherwise	840	0.17	0.38	459	0.17	0.38	0.00	0.301	

Source: Authors' calculations based on student surveys.

**Significant at the .01 level. * Significant at the .05 level.

The group difference was evaluated applying the same statistical model used to estimate the impact, with only the assignment group and blocking variables as controls. The size of the analytic sample is 1408 (911 in intervention and 497 in control). Equivalence tests were conducted using non-missing observations of baseline variables. In the impact estimation, the dummy variable adjustment method was used for missing baseline variables as covariates.

For race/ethnic indicators used as covariates in the study, we counted a student in a particular race/ethnic group if s/he identified himself/herself as that racial/ethnic group, regardless of whether the student identified with one or more other race/ethnic groups. In addition, we also examined the racial/ethnic distribution based on the conventional U.S. Census definition of race/ethnic groups, where a group is defined excluding mixed race individuals. These mutual exclusive race/ethic categories are presented in the bottom rows of the table.

Exhibit C.18: Summary Statistics of Key Baseline Measures for Student Characteristics Analytic Sample for Estimating Impact on Initiation of Sexual Activity, Excluding Two Blocks that Included Withdrawn Schools (31 schools)

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Initiation of sexual activity (1=ever had sex, 0=otherwise)	779	0.06	0.24	404	0.06	0.25	0.00	0.544	X
Age in years	893	12.23	0.62	482	12.24	0.58	-0.01	0.645	X
Grade level at start of study ("1"=7th grade, "0"=otherwise/8th grade)	901	0.95	0.22	491	0.89	0.32	0.06	0.423	
Female ("1"=female, "0"=Otherwise)	897	0.53	0.50	486	0.54	0.50	-0.02	0.712	X
Bisexual/homosexual ("1"=Bi/homosexual; "0"=Otherwise)	887	0.02	0.16	480	0.02	0.14	0.01	0.610	
Student using non-English at home ("1"=Use non English at home; "0"=Otherwise)	908	0.26	0.44	494	0.38	0.49	-0.12	0.000	X
Parent does not speak English ("1"=Does not speak English; "0"=Otherwise)	899	0.02	0.12	490	0.02	0.15	-0.01	0.055	
Receiving mostly As and Bs for grade ("1"=As and Bs, "0"=Otherwise)	854	0.72	0.45	461	0.71	0.45	0.01	0.819	X
Native Hawaiian ("1"=Native Hawaiian including mixed race, "0"=Otherwise)	896	0.43	0.50	490	0.49	0.50	-0.06	0.422	X

Characteristic	Intervention N.	Intervention Mean	Intervention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Asian ("1"=Asian including mixed race, "0"=Otherwise)	896	0.73	0.45	490	0.75	0.43	-0.02	0.068	X
Mutually exclusive race/ ethnicity classifications:									
Non-Hispanic Hawaiian ("1"=Native Hawaiian including no mixed race, "0"=Otherwise)	834	0.02	0.14	455	0.03	0.17	-0.01	0.241	
Non-Hispanic Asian ("1"Asian including no mixed race, "0"=Otherwise)	834	0.20	0.40	455	0.24	0.42	-0.04	0.179	
Non-Hispanic Black (Includes no other race)	834	0.00	0.06	455	0.00	0.07	0.00	0.912	
Non-Hispanic White ("1"=White including mixed race, "0"=Otherwise)	834	0.07	0.25	455	0.06	0.24	0.01	0.283	
Non-Hispanic with two or more races	834	0.49	0.50	455	0.45	0.50	0.04	0.327	
Hispanic ("1"=Hispanic, "0"=Otherwise	837	0.17	0.38	457	0.17	0.38	0.00	0.315	

Source: Authors' calculations based on student surveys.

**Significant at the .01 level. * Significant at the .05 level.

The group difference was evaluated applying the same statistical model used to estimate the impact, with only the assignment group and blocking variables as controls. The size of the analytic sample is 1402 (908 in intervention and 494 in control). Equivalence tests were conducted using non-missing observations of baseline variables. In the impact estimation, the dummy variable adjustment method was used for missing baseline variables as covariates.

For race/ethnic indicators used as covariates in the study, we counted a student in a particular race/ethnic group if s/he identified himself/herself as that racial/ethnic group, regardless of whether the student identified with one or more other race/ethnic groups. In addition, we also examined the racial/ethnic distribution based on the conventional U.S. Census definition of race/ethnic groups, where a group is defined excluding mixed race individuals. These mutual exclusive race/ethic categories are presented in the bottom rows of the table.

Exhibit C.19: Summary Statistics of Key Baseline Measures for Student Characteristics Analytic Sample for Estimating Impact on Knowledge of Pregnancy and STI Prevention, Excluding Two Blocks that Included Withdrawn Schools (31 schools)

Characteristic	Inter-vention N.	Inter-vention Mean	Inter-vention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Knowledge of pregnancy and STI prevention (proportion of correct responses to 10 questions in baseline survey)	929	0.64	0.20	491	0.61	0.22	0.03	0.216	X
Age in years	928	12.23	0.62	501	12.24	0.58	-0.02	0.130	X
Grade level at start of study ("1"=7th grade, "0"=otherwise/8th grade)	938	0.95	0.22	510	0.88	0.32	0.07	0.128	
Female ("1"=female, "0"=Otherwise)	934	0.52	0.50	505	0.54	0.50	-0.02	0.854	X
Bisexual/homosexual ("1"=Bi/homosexual; "0"=Otherwise)	922	0.02	0.16	498	0.02	0.15	0.00	0.707	
Student using non-English at home ("1"=Use non English at home; "0"=Otherwise)	944	0.26	0.44	513	0.38	0.49	-0.12	0.001	X
Parent does not speak English ("1"=Does not speak English; "0"=Otherwise)	936	0.02	0.13	509	0.02	0.15	-0.01	0.337	
Receiving mostly As and Bs for grade ("1"=As and Bs, "0"=Otherwise)	890	0.73	0.45	478	0.71	0.46	0.02	0.578	X
Native Hawaiian ("1"=Native Hawaiian including mixed race, "0"=Otherwise)	932	0.43	0.50	509	0.49	0.50	-0.06	0.169	X

Characteristic	Inter-vention N.	Inter-vention Mean	Inter-vention SD	Control N.	Control Mean	Control SD	Group Difference in Mean	*p*-value for Difference	Included as covariates in benchmark estimation model
Asian ("1"=Asian including mixed race, "0"=Otherwise)	932	0.73	0.44	509	0.75	0.43	-0.02	0.524	X
Mutually exclusive race/ethnicity classifications:									
Non-Hispanic Hawaiian ("1"=Native Hawaiian including no mixed race, "0"=Otherwise)	867	0.02	0.13	472	0.03	0.17	-0.01	0.198	
Non-Hispanic Asian ("1"Asian including no mixed race, "0"=Otherwise)	867	0.20	0.40	472	0.24	0.42	-0.04	0.496	
Non-Hispanic Black (Includes no other race)	867	0.00	0.06	472	0.00	0.07	0.00	0.913	
Non-Hispanic White ("1"=White including mixed race, "0"=Otherwise)	867	0.07	0.25	472	0.06	0.24	0.01	0.416	
Non-Hispanic with two or more races	867	0.49	0.50	472	0.46	0.50	0.04	0.913	
Hispanic ("1"=Hispanic, "0"=Otherwise	870	0.17	0.38	474	0.17	0.38	0.00	0.912	

Source: Authors' calculations based on student surveys.

**Significant at the .01 level. * Significant at the .05 level.

The group difference was evaluated applying the same statistical model used to estimate the impact, with only the assignment group and blocking variables as controls. The size of the analytic sample is 1457 (944 in intervention and 513 in control). Equivalence tests were conducted using non-missing observations of baseline variables. In the impact estimation, the dummy variable adjustment method was used for missing baseline variables as covariates.

For race/ethnic indicators used as covariates in the study, we counted a student in a particular race/ethnic group if s/he identified himself/herself as that racial/ethnic group, regardless of whether the student identified with one or more other race/ethnic groups. In addition, we also examined the racial/ethnic distribution based on the conventional U.S. Census definition of race/ethnic groups, where a group is defined excluding mixed race individuals. These mutual exclusive race/ethic categories are presented in the bottom rows of the table.]

Exhibit C.20: Summary Statistics of Key Baseline Measures for Student and School Characteristics
Alternative Analytic Sample, Excluding Two Blocks that Included Withdrawn Schools
Baseline School Characteristics
All Outcome Measures

School Characteristics	Total* N	Total Mean	Total SD	Treat-ment** N	Treatment Mean	Treat-ment SD	Control *** N	Control Mean	Control SD	Difference Mean	*p*-value for Difference
Classroom size (number of students)	30	14.43	2.97	16	15.09	3.02	14	13.67	2.83	1.42	0.196
School size (total enrollment)	30	518	371	16	604	371	14	419	359	184.21	0.179
Poverty level (ratio of students eligible for free or reduced-price lunch)	30	0.58	0.17	16	0.57	0.18	14	0.59	0.15	-0.03	0.647
English proficiency (ratio of students classified as English language learners)	29	0.07	0.08	15	0.06	0.05	14	0.07	0.10	-0.01	0.744
Performance (1= not meeting AYP, in restructuring, or a Race to the Top priority school, 0=otherwise)	30	0.07	0.25	16	0.13	0.34	14	0.00	0.00	0.13	0.171

* Total: N=31
** Treatment: N=16
*** Control: N=15
Source: Hawai'i Department of Education and Common Core Data (2010-11 and 2011-12), National Center for Education Statistics.

Exhibit C.21: Covariates Used in Impact Estimation

			Alternative	Specifications	
Covariate	**Covariate Description**	**Unadjusted**	**Adjusted for Baseline Outcome**	**Adjusted for All Covariates**	**Adjusted for Select Covariates (reported model)**
Study design variables					
Treatment indicator	A dummy variable indicating whether the school is offering the Pono Choices curriculum (Impact measure)	X	X	X	X
Blocking indicators	A set of dummy variables indicating the school is in a given block. The schools are blocked based on island, school type (regular public vs. public charter/private), timing of planned sexual health instruction, and recruitment semester.	X	X	X	X
Semester cohort indicators	Indicators for the semester in which the student is offered the programming			X	X
School characteristics					
Classroom size	Average student-to-teacher ratio			X	
School size	Total student enrollment			X	
Poverty level	Percentage of students eligible for free/reduced price lunch			X	X
English proficiency	Percentage of students classified as English language learners			X	
Performance	1= school does not meeting AYP, is in restructuring, or is a priority school in Race to the Top			X	
Student characteristics					
Outcome at baseline	Outcome measured at baseline		X	X	X
Age	Age in years			X	X
Gender	Indicator for female			X	X
Sexual orientation	Indicator for bisexual or homosexual orientation			X	

			Alternative	Specifications	
Covariate	Covariate Description	Unadjusted	Adjusted for Baseline Outcome	Adjusted for All Covariates	Adjusted for Select Covariates (reported model)
Non-English speaking at home	Indicator for the student not speaking English at home			X	X
Non-English speaking parents	Indicator for parent(s) not speaking English			X	
Race/ethnicity: Native Hawaiian	Indicator for the student identifying as Native Hawaiian			X	X
Race/ethnicity: Asian	Indicator for the student identifying as Asian			X	
Race/ethnicity: White	Indicator for the student identifying as White			X	
Race/ethnicity: Hispanic	Indicator for the student identifying as Hispanic			X	
Academic grade received	Indicator for the student receiving mostly A's and B's			X	X
Grade	Indicator for the student being in 7th grade at baseline			X	

Exhibit C.22: Implementation Evaluation Methods

Implementation Element	Methods Used to Address Each Implementation Element
Adherence	
How often were sessions offered? How many were offered?	The total number of sessions documented in the attendance logs. Average session duration calculated as the average of the observed session lengths, measured in minutes. Average weekly frequency calculated as the total number of sessions divided by the total number of weeks between the first and last sessions of the curriculum.
What and how much was received?	Average of number of sessions attended calculated as the average of the number of sessions that students attended. Attendance data is also used to calculate the percentage of students attending at least 75% of the sessions. Percentage of sessions attended is calculated as the total number of sessions attended divided by the total number of sessions offered.
What content was delivered to youth?	The topics covered are defined by the individual modules and activities within each module. Facilitator logs provide data on activities completed for 100% of the sessions. Observers' logs for 22% of the sessions show a 98% agreement with Facilitator logs. Frequency data is used to identify any specific topics (activities within sessions) that might have been skipped more frequently than others.
Who delivered material to youth?	Intervention staff interviewed control school teachers at the end of each semester and gathered information about who delivered the sexual health curriculum in those classes. (In intervention schools the curriculum was delivered by regular health or physical education teachers.)
Quality	
Quality of staff-participant interactions	Calculated as the percentage of observed interactions where the independent evaluator scored the interaction as "4" or "5" on each of the items on the observation form with 5 being "Excellent".
Quality of youth engagement with program	Calculated as the percentage of sessions where the independent evaluator scored "How actively did students participate in discussions and activities?" as a "4" or "5." On a scale of 1 to 5 with 5 being "Excellent".
Counterfactual	
Experiences of counterfactual condition	The data from exit interviews with control teachers on experiences of the counterfactual are presented as frequency counts and percentages.

Implementation Element	Methods Used to Address Each Implementation Element
Context	
Other TPP programming available or offered to study participants (both intervention and counterfactual)	Service provision was monitored by both implementation staff and evaluators as part of the fidelity monitoring and maintenance during implementation.
Substantial unplanned adaptation(s)	Adaptations were captured on observer fidelity logs.
Counterfactual	
Experiences of counterfactual condition	The data from exit interviews with control teachers on experiences of the counterfactual are presented as frequency counts and percentages.
Context	
Other TPP programming available or offered to study participants (both intervention and counterfactual)	Service provision was monitored by both implementation staff and evaluators as part of the fidelity monitoring and maintenance during implementation.
External events affecting implementation	The evaluation team discussed current events implementation team members to note any events that seemed likely to affect implementation.
Substantial unplanned adaptation(s)	Adaptations were captured on observer fidelity logs.

APPENDIX D. RECRUITMENT OF SCHOOLS

School recruitment was conducted by the study Implementation Team. (The external evaluators did not participate in recruitment.) Recruitment focused on the 83 middle schools (54 regular public schools and 29 charter schools) in the State of Hawai'i. Recruitment efforts were conducted twice, resulting in recruiting two cohorts of schools. Recruiting materials included an HIDOE approval letter, the Pono Choices brochure or information sheet, a description of and invitation into the project, a sample MOA/MOU, and a follow-up script or checklist used when speaking to principals and schools.

Recruitment initially focused on the 73 regular public and charter middle schools on the islands of O'ahu, Hawai'i and Maui, since these are the 3 islands where the majority of schools and students are located. With the aim of increasing the number of study schools, in a second recruitment we expanded statewide to include the islands of Lāna'i, Moloka'i, and Kaua'i. This resulted in a total of 83 schools being actively recruited or invited into the study.

The first recruitment effort began in July 2011 with an email to the principals of 73 schools that described the project and invited their school to participate. This email was followed by a letter to complex area (sub-district) superintendents and a follow-up letter and/or phone call to principals. These efforts resulted in 25 inquiries or interested schools and 19 schools from two islands (Cohort 1) signing an MOA/MOU and committing to random assignment for Spring 2012 implementation. Three other schools committed to delayed implementation and were included in the next randomization.

Since the first recruitment effort resulted in fewer than the target of 30 schools, a second recruitment was conducted in the Spring of 2012. Discussions with HIDOE resulted in excluding middle schools that included 6th grade because they were being considered for another OAH study.

After excluding these schools, the remaining 34 regular public middle or public charter schools that were not already in the project were invited into the project through letters and emails sent to principals. These invitations were followed by phone calls and resulted in 20 new schools inquiring or showing interest and 14 new schools signing an MOA/MOU and committing to random assignment for the 2012-2013 school year. These and the 3 schools secured from the previous recruitment effort became the 17 schools in Cohort 2.

During the second recruitment effort a private school heard about the study and asked to participate. The one exclusion criteria was public middle schools that included 6th grade because these schools were being considered for another OAH study. The private school did not meet this exclusion criterion and did meet the inclusion criterion of covering sexual health in middle school. Therefore, we did not have a reason to exclude the school from being in the sample and included them when the second cohort of schools was randomly assigned.

Overall, 36 schools were recruited and randomized. These schools represent 21 out of the 54 (39%) regular public middle schools, 14 out of the 29 (48%) public charter schools in the state, and one private school. Thirteen out of 15 complex areas (87%) or sub-districts are represented in the project, and where there is more than one study school in a complex area, they include at least one intervention school and one control school.

APPENDIX E. ESTIMATION MODEL

The program impacts were estimated as the differences in the student outcome measures between program and control groups 1-year after baseline data collection, after adjusting for the stratification imposed by design and variables measured at baseline. The impact was estimated as intent-to-treat effects of the intervention, including all random-assigned schools and study-eligible cohorts of students in the analysis sample, regardless of the level of actual participation in the intervention.

To account for the nested nature of the data, the study used a mixed-level model for the estimation of the program impact. The model is specified as a two-level random-intercept model, in which the student (first) level is nested in the school (second) level. For student *i* and school *j*, for $i = 1 \ldots N$ and $j = 1 \ldots K$, the model is specified as the following system of equations:

$$\text{(Eq. 1)} \quad Y_{ij} = \alpha_j = \sum_{q=1}^{Q} \beta_q X_{qij} + \varepsilon_{ij}$$

$$\text{(Eq. 2)} \quad \alpha_j = \gamma_0 + \gamma_1 \left(Treatment_j \right) + \sum_{s=2}^{S} \gamma_s W_{sj} + u_j$$

Where Y_{ij} denotes a student outcome, *Treatment* is a dummy variable indicating whether school *j* is randomly assigned to receive Pono Choices (*Treatment* = "1") or not (Treatment = "0"), and its coefficient γ_1 represents the estimated effects of Pono Choices on the student outcome. X_q is a student-level covariate at baseline; W_{sj} represents a school-level covariate (a characteristic or blocking variable) at baseline; and β_q and γ_s are estimators for marginal effects of individual- and school-level covariates, respectively. The model assumes two random error terms: ε_{ij} is the error term specific to student *i* in school *j* and u_j is the error term specific to the *j*-th school, representing the random school effects.

Substituting the school-level equation into the student-level equation above, the system of equations are rewritten as:

(Eq. 3) $$Y_{ij} = \gamma_0 + \gamma_1\left(Treatment_j\right) + \sum_{q=1}^{Q} \beta_q X_s W_{sj} + u_j + \varepsilon_{ij}$$

The study estimated this reduced-form model (Eq. 3). When the outcome is binary (e.g., behavioral outcomes for confirmatory analyses), the model is estimated assuming a logistic distribution for ε_{ij}. When the outcome is continuous (e.g., non-behavioral outcomes for exploratory analyses), the model is estimated assuming a normal distribution.

To answer each research question, we tested the null hypothesis that there was no difference between the groups $\left(H_0 : \gamma_1 = 0\right)$. If the null hypothesis was rejected by a two-tailed test at the 5 percent significance level, we concluded that the outcome was different for students who participated in classes offering Pono Choices and students who did not.

APPENDIX F. SUPPLEMENTAL EXHIBITS FOR IMPLEMENTATION FINDINGS FOR CHAPTER IV

Exhibit F.1: Summary of Implementation Findings

Implementation Measures	Spring 2012 Semester Cohort	Fall 2012 Semester Cohort	Spring 2013 Semester Cohort	TOTAL
Fidelity				
Percent of activities completed	98%	94%	98%	98%
Dosage				
Mean percent of program content received	94%	95%	93%	94%
Percent of students who received at least 75% of content	92%	95%	93%	94%
Overall Quality				
Average quality rating (on 5 point scale from "1" low to "5" high)	4.33	4.28	4.26	4.27
Percent of curriculum modules with overall quality score of 4.0 or higher	73%	83%	76%	77%
Student Engagement				
Average student engagement rating (5 point scale from "1" low to "5" high)	4.39	4.47	4.61	4.54
Percent of modules with student engagement score of 4.0 or higher	85%	84%	87%	86%

Source: Observers' Lesson Delivery Logs, Observer's Program Observation Form, Attendance logs

Exhibit F.2: Teacher-Reported Activities Completed across All Classes/Sections by Module

Module	Percent of Scheduled Activities Completed Mean	Percent of Scheduled Activities Completed Median	Number of Activities Completed Across All Classes/Sections Sum	Number of Activities Scheduled Across All Classes/Sections Sum
Module 1: Introduction	100%	100%	230	230
Module 2: Making Responsible Choices about Sex	98%	100%	315	322
Module 3: Reproductive Anatomy & Puberty	94%	100%	391	414
Module 4: Communication & Healthy Relationships	94%	100%	296	315
Module 5: Refusal Skills	97%	100%	340	352
Module 6: Pregnancy & Birth Control	97%	100%	341	352
Module 7: Understanding STIs	99%	100%	364	368
Module 8: Preventing STIs	97%	100%	357	368
Module 9: Negotiation and Refusal Skills	93%	98%	307	368
Module 10: Review & Empowerment	97%	100%	269	276
OVERALL	**98%**	**100%**	**3,210**	**3,365**

Source: Facilitators' Lesson Delivery Logs

Exhibit F.3: Observers' Assessment of Activities Completed

Module	Number of Observations	Activities Scheduled	Activities Completed	Activities Completed with Changes	Activities Completed with Changes	Percent Completed	Percent Completed
	Count	Sum	Sum	Sum	Percent	Mean	Median
Module 1: Introduction	21	127	126	24	19%	100%	100%
Module 2: Making Responsible Choices about Sex	10	56	50	11	22%	89%	100%
Module 3: Reproductive Anatomy & Puberty	9	77	76	33	37%	99%	100%
Module 4: Communication & Healthy Relationships	15	99	95	39	27%	96%	100%
Module 5: Refusal Skills	9	64	64	22	34%	100%	100%
Module 6: Pregnancy & Birth Control	12	91	90	22	24%	99%	100%
Module 7: Understanding STIs	10	73	72	26	36%	99%	100%
Module 8: Preventing STIs	16	115	115	46	31%	100%	100%
Module 9: Negotiation and Refusal Skills	11	82	80	17	21%	98%	100%
Module 10: Review & Empowerment	15	86	85	46	31%	98%	100%
OVERALL	**128**	**870**	**853**	**238**	**28%**	**98.4%**	**100%**

Source: Observers' Lesson Delivery Logs.

Note: The evaluators reviewed 128 observations/lesson delivery logs conducted by the external evaluation team, as well as 605 lesson delivery logs completed by program teachers. Overall fidelity to the curriculum as intended was very high and the observed challenges and adaptations were only minor. The most frequent "change" to the program was spending more time on specific components than planned. Neither the observations nor the teachers' logs raised any significant fidelity concerns or deviations from implementation as intended. However, the evaluators provided the implementation team with a summary of the general challenges experienced and the adaptations made to the program by teachers, along with module-specific challenges, for use in refining implementation guidance for disseminating the curriculum in the future.

Exhibit F.4: Comparison of Teachers' and Observers' Assessment of Activities Delivered by Module

Module	Teachers' Reported Total Activities Completed	Observers Reported Total Number of Activities Completed	Concurrence Between Teachers and Observers in Number of Activities
Module 1: Introduction	120	120	100%
Module 2: Making Responsible Choices about Sex	49	50	98%
Module 3: Reproductive Anatomy & Puberty	86	86	100%
Module 4: Communication & Healthy Relationships	87	88	99%
Module 5: Refusal Skills	64	66	97%
Module 6: Pregnancy & Birth Control	86	87	99%
Module 7: Understanding STIs	77	76	99%
Module 8: Preventing STIs	105	108	97%
Module 9: Negotiation and Refusal Skills	79	88	90%
Module 10: Review & Empowerment	87	88	99%
TOTAL	**840**	**857**	**98%**

Source: Facilitators' Lesson Delivery Logs and Observers' Lesson Delivery Logs for the 128 lessons observed by evaluators.

Exhibit F.5: Student Attendance by Module

Module	Did not Attend Count	Did not Attend Row %	Attended Count	Attended Row %
Module 1: Introduction	61	5.4%	1,069	94.6%
Module 2: Making Responsible Choices about Sex	50	4.4%	1,080	95.6%
Module 3: Reproductive Anatomy & Puberty	49	4.3%	1,081	95.7%
Module 4: Communication & Healthy Relationships	72	6.4%	1,058	93.6%
Module 5: Refusal Skills	55	4.9%	1,075	95.1%
Module 6: Pregnancy & Birth Control	92	8.1%	1,038	91.9%
Module 7: Understanding STIs	74	6.5%	1,056	93.5%
Module 8: Preventing STIs	69	6.1%	1,061	93.9%
Module 9: Negotiation and Refusal Skills	89	7.9%	1,041	92.1%
Module 10: Review & Empowerment	66	5.8%	1,064	94.2%
ACROSS ALL MODULES		**1,130**	**Mean =**	**94%**

Source: Attendance Logs.

Note: N=1,130. While a total of 1,158 treatment school students had parent consent to participate, attendance data was not reported for 28 students, including 7 students who chose not to participate, and 13 students who were indicated by their teachers as having moved or withdrawn from the class.

Exhibit F.6.1: Time Spent on Each Module – Cohort 1 – Spring 2012

Module	Number of Minutes Spent Mean	Number of Minutes Spent Median	Number of Minutes Spent Maximum	Number of Minutes Intended	Difference Between Intended and Mean
Module 1: Introduction	59	59	71	60	1
Module 2: Making Responsible Choices about Sex	58	59	60	60	2
Module 3: Reproductive Anatomy & Puberty	70	70	74	60	-10
Module 4: Communication & Healthy Relationships	62	64	71	60	-2
Module 5: Refusal Skills	76	76	91	60	-16
Module 6: Pregnancy & Birth Control	69	64	85	60	-9
Module 7: Understanding STIs	57	56	58	60	3
Module 8: Preventing STIs	71	67	88	60	-11
Module 9: Negotiation and Refusal Skills	77	74	98	60	-17
Module 10: Review & Empowerment	64	67	95	60	-4

Exhibit F.6.2: Time Spent on Each Module – Cohorts 2 & 3 – Fall 2012 – Spring 2013

Module	Number of Minutes Spent Mean	Number of Minutes Spent Median	Number of Minutes Spent Maximum	Number of Minutes Intended	Difference Between Intended and Mean
Module 1: Introduction	22	21	28	30	8
Module 2: Making Responsible Choices about Sex	40	36	75	60	20
Module 3: Reproductive Anatomy & Puberty	46	56	62	60	14
Module 4: Communication & Healthy Relationships	52	56	71	60	8
Module 5: Refusal Skills	39	32	76	60	21
Module 6: Pregnancy & Birth Control	46	49	64	60	14
Module 7: Understanding STIs	44	44	67	60	16
Module 8: Preventing STIs	45	46	83	60	15
Module 9: Negotiation and Refusal Skills	46	49	62	60	14
Module 10: Review & Empowerment	49	51	64	60	11

Exhibit F.7: Number and Percentage of Modules with Average Overall Quality Score of 4.0 or More by Module

Module	Number of Modules with Average Quality Score of 4.0 or Higher	Total Number of Each Module Observed Across All Schools	Average Rating	Percent of Modules with Average Quality Score of 4.0 or Higher
Module 1: Introduction	18	21	4.50	86%
Module 2: Making Responsible Choices about Sex	8	10	4.14	80%
Module 3: Reproductive Anatomy & Puberty	8	9	4.53	89%
Module 4: Communication & Healthy Relationships	10	14	4.23	71%
Module 5: Refusal Skills	7	10	3.88	70%
Module 6: Pregnancy & Birth Control	7	12	4.20	58%
Module 7:| Understanding STIs	7	10	4.29	70%
Module 8: Preventing STIs	13	16	4.45	81%
Module 9: Negotiation and Refusal Skills	9	11	4.21	82%
Module 10: Review & Empowerment	11	15	4.35	73%
OVERALL	**98**	**128**	**4.27**	**77%**

Source: Observers' Program Observation Forms.

Exhibit F.8: Student Engagement Scores by Module

Module	Average Rating	Percent of Modules with Average Quality Score of 4.0 or Higher
Module 1: Introduction	4.62	81%
Module 2: Making Responsible Choices about Sex	4.60	90%
Module 3: Reproductive Anatomy & Puberty	4.44	78%
Module 4: Communication & Healthy Relationships	4.43	100%
Module 5: Refusal Skills	4.00	60%
Module 6: Pregnancy & Birth Control	4.33	83%
Module 7: Understanding STIs	4.10	70%
Module 8: Preventing STIs	4.56	94%
Module 9: Negotiation and Refusal Skills	4.55	82%
Module 10: Review & Empowerment	4.53	93%
OVERALL	**4.54**	**86%**

Source: Observers' Program Observation Forms.

Exhibit F.9: Number and Percentage of Modules with Score of 4.0 or More by Observation Item

Observation Item	Number of Items with Avg. Quality Score ≥ 4.0	Total Number of Observed Modules	Average Rating	Percent of Items with Avg. Quality Score ≥ 4.0
1. How clear were teachers' explanations?	120	128	4.56	94%
2. To what extent did teacher keep track of time?	110	128	4.23	86%
3. To what extent was presentation of material rushed?	93	128	4.09	73%
4. To what extent did participants understand material?	115	128	4.52	90%
5. How actively did students participate in discussion or activities?	108	128	4.45	84%
6a. Teacher knowledge of program content.	117	128	4.20	91%
6b. Teacher level of enthusiasm.	102	128	4.09	80%
6c. Teacher poise and confidence.	117	128	4.46	91%
6d. Teacher rapport and communication with participants.	108	128	4.30	84%
6e. Teacher effectively addressed questions/ concerns.	116	128	4.37	91%
7. Overall quality of program session.	102	128	4.06	80%
OVERALL	**1208**	**1408**	**4.30**	**85.8%**

Source: Observers' Program Observation Forms.

Exhibit F.10: Curriculum Components in Control Schools

Key Curriculum Components	Number of Schools Spring 2012	Number of Schools Fall 2012	Number of Schools Spring 2013	Total Number of Control Schools *	Percent of All Control Schools	Percent of All Control School Students Receiving Each Component
Reproductive Anatomy	4	4	8	9	52.9%	64.3%
Pregnancy prevention	6	7	10	13	76.5%	87.7%
STI prevention	5	7	9	12	70.6%	81.2%
Refusal skills	4	5	9	10	58.8%	66.8%
Condom demonstration	0	3	3	5	29.4%	23.7%
TOTAL	7	8	13	17	100%	

Source: Teacher interviews

Note: * Total number of schools does not equal the sum of schools participating across the three semesters because schools varied in the number of semesters they provided health instruction during the study period.

Exhibit F.11: Instructor Characteristics in Intervention vs. Control Schools

	Program School Instructors	Program School Instructors	Control School Instructors	Control School Instructors
Teacher Experience	**N=18**	**%**	**N=16**	**%**
Curriculum Delivered by:				
Regular Classroom Teacher	18	100%	6	37.5%
Guest Speaker (Sexual Health Specialist)	0	0	2	12.5%
Both	0	0	8	50.0%
Prior Experience Teaching Sexual Health:				
0 years	6	33.3%	5	31.2%
1-4 years	4	22.2%	4	25.0%
5 or more years	8	44.4%	6	37.5%

Source: Teacher interviews

APPENDIX G. SUPPLEMENTAL EXHIBITS FOR IMPACT FINDINGS IN CHAPTER IV

Exhibit G.1: Summary of Outcome Measures at 1-Year Follow-Up

Variable	Total Mean/ Percent	Total Std. Dev	Total Obs	Interv. Group Mean/ Percent	Interv. Group Std. Dev	Interv. Group Std. Obs	Control Group Mean/ Percent	Control Group Std. Dev	Obs	Min. Value	Max. Value
Behavioral outcomes											
Engagement in high-risk sexual behavior(binary indicator)	0.018	0.133	1,494	0.014	0.116	961	0.026	0.160	533	0	1
Initiation of sexual intercourse (binary indicator)	0.099	0.298	1,488	0.094	0.292	958	0.108	0.310	530	0	1
Non-behavioral outcomes											
Knowledge of pregnancy and STI prevention (percent correct responses)	0.663	0.209	1,546	0.719	0.185	995	0.561	0.212	551	0	1
Attitudes toward healthy sexual behaviors (average score on scale of 1-4 where 4=very important)	3.599	0.499	1,537	3.611	0.483	990	3.578	0.526	547	1	4
Skills in managing relationships and choices (average score on scale of 1-4 where 1=very difficult and 4=very easy)	2.981	0.600	1,425	2.991	0.568	943	2.961	0.658	482	1	4
Intention to have sex (average score on a scale of 1-4 where 4=Very likely)	3.336	0.896	1,374	3.334	0.900	902	3.339	0.888	472	1	4
Intention to use condom during intercourse (average score on a scale of 1-4 where 4=Very likely)	3.700	0.689	1,361	3.696	0.694	911	3.709	0.679	450	1	4
Intention to use effective birth control (average score on a scale of 1-4 where 4=Very likely)	3.444	0.863	1,196	3.464	0.852	823	3.399	0.885	373	1	4

Exhibit G.2: A Summary of Behavioral Outcome Measures, Alternative Sample Excluding the Blocks with Withdrawn Schools

Variable	Total Mean/ Percent	Total Std. Dev	Total Obs	Treat-ment Group Mean/ Percent	Treat-ment Group Std. Dev	Treat-ment Group Obs	Control Group Mean/ Percent	Control Group Std. Dev	Control Group Obs	Min. Value	Max. Value
Engagement in high-risk sexual behavior (binary indicator)	0.016	0.127	1,408	0.013	0.114	911	0.022	0.147	497	0	1
Initiation of sexual intercourse (binary indicator)	0.100	0.300	1,402	0.095	0.293	908	0.109	0.312	494	0	1

Source: IMPAQ staff calculations based on student surveys.

Exhibit G.3: Estimated Impacts on Engagement in High-Risk Sexual Behavior at 1-Year Follow-up

Odds of Having Engaged in Unprotected Sex Ratio	Odds of Having Engaged in Unprotected Sex S.E.	Treatment Group Probability	S.E.	Control Group Probability	S.E.	Difference in Probability	S.E.	*p*-value of Impact (odds not equal to 1)	N
0.666	0.3411	0.015	0.0045	0.022	0.0068	-0.007	0.0090	0.428	1,494

Source: IMPAQ staff calculations based on student surveys.
Note: Predicted probability (estimated percentage) of engaging in unprotected sex.

Exhibit G.4: Estimated Impacts on Initiation of Sexual Activity at 1-Year Follow-up

Odds of Having Initiated Sexual Activity Ratio	Odds of Having Initiated Sexual Activity S.E.	Treatment Group Probability	S.E.	Control Group Probability	S.E.	Difference in Probability	S.E.	*p*-value of Impact (odds not equal to 1)	N
0.984	0.2228	0.098	0.0097	0.100	0.0127	-0.001	0.0171	0.944	1,488

Source: IMPAQ staff calculations based on student surveys.
Note: Predicted probability (estimated percentage) of having initiated sexual activity.

Exhibit G.5: Estimated Impacts on Non-Behavioral Outcomes

Outcome	Treatment Group Mean	Control Group Mean	Difference in Means	S.E.	p-Value of Difference	Effect Size (Hedges's g)	N
Knowledge of pregnancy and STI prevention (proportion of correction responses to 10 questions related to TPP/STI prevention)	0.718	0.564	0.154**	0.0162	<0.001	0.788	1,546
Attitudes toward healthy sexual behaviors (average score on scale of 1-4 where 4=very important)	3.602	3.591	0.011	0.0294	0.708	0.022	1,537
Skills in managing relationships and choices (average score on scale of 1-4 where 1=very difficult and 4=very easy)	2.986	2.972	0.015	0.0395	0.709	0.025	1,425
Intention to have sex (average score on a scale of 1-4 where 4=Very likely)	3.343	3.321	0.02130	0.0960	0.824	0.024	1,374
Intention to use condom during intercourse (average score on a scale of 1-4 where 4=Very likely)	3.686	3.701	-0.01435	0.0612	0.815	-0.021	1,361
Intention to use effective birth control (average score on a scale of 1-4 where 4=Very likely)	3.438	3.395	0.04337	0.0843	0.607	0.050	1,196

Source: IMPAQ staff calculations based on student surveys.

**Significant at the .01 level.

Note: The knowledge measure is statistically significant after adjusting for multiple comparisons. The measure remains statistically significant using either the Benjamini & Hochberg Method or the Bonferroni method to adjust for six comparisons under the non-behavioral outcome domain. The measure also remained statistically significant using either method to adjust for eight comparisons across behavioral and non-behavioral outcome

Exhibit G.6: Alternative Estimates of Program Effects on Probability of Engagement in High-Risk Sexual Behavior in the Last Three Months, at 1-Year Follow-up

	N	Difference in Probability (Treatment – Control)	S.E.	Odds-Ratio	S.E.	*p*-value of Impact (odds not equal to 1)
Benchmark model presented in Table 4						
Adjusted for select covariates, with imputed covariates, estimated with random effects logit	1,494	-0.0069	0.0090	0.666	0.3411	0.428
Alternative covariate specifications						
Unadjusted (no covariate adjustment)	1,494	-0.014	0.0088	0.487	0.2093	0.094
Adjusted for baseline outcome measure	1,494	-0.014	0.0090	0.469	0.2063	0.085
Adjusted for all covariates	1,494	-0.010	0.0134	0.565	0.4207	0.443
Alternative missing data treatment						
Adjusted for select covariates, with no listwise deletion.	1,177	0.016	0.0159	4.276	5.5872	0.266
Alternative estimation method						
Adjusted for select covariates, with imputed covariates, estimated with random effects probit.	1,494	-0.0078	0.0088	--	--	0.364
Alternative study sample						
Exclude two blocks that included schools that withdrew immediately after the random assignment.	1,408	-0.0069	0.0087	0.646	0.3379	0.403

Source: IMPAQ staff calculations based on student surveys.

Exhibit G.7: Alternative Estimates of Program Effects on Probability of Initiation of Sexual Activity, at 1-Year Follow-up

	N	Difference in Probability (Treatment – Control)	S.E.	Odds-Ratio	S.E.	*p*-value of Impact (odds not equal to 1)
Benchmark model presented in Table 4						
Adjusted for select covariates, with imputed covariates, estimated with random effects logit	1,488	-0.001	0.0171	0.993	0.2244	0.974
Alternative covariate specifications						
Unadjusted (no covariate adjustment)	1,488	-0.007	0.0183	0.923	0.1892	0.696
Adjusted for baseline outcome measure	1,488	-0.007	0.0170	0.914	0.1962	0.674
Adjusted for all covariates	1,488	-0.005	0.0230	0.938	0.2886	0.835
Alternative missing data treatment						
Adjusted for select covariates, with no listwise deletion.	1,153	0.012	0.0175	1.210	0.3505	0.508
Alternative estimation method						
Adjusted for select covariates, with imputed covariates, estimated with random effects probit.	1,488	0.002	0.0170	---	---	0.922
Alternative study sample						
Exclude two blocks that included schools that withdrew immediately after the random assignment.	1,402	-0.003	0.0172	0.9658	0.2191	0.878

Source: IMPAQ staff calculations based on student surveys.

Exhibit G.8: Alternative Estimates of Program Effects on Knowledge Measure, at 1-Year Follow-up

	N.	Difference (Treatment – Control)	S.E.	*p*-value	Effect size (Hedges's g)
Benchmark model presented in Exhibit G.5.					
Benchmark: Adjusted for select covariates, with imputed covariates, estimated with REML	1,546	0.151	0.0159	<0.001	0.730
Alternative covariate specifications					
Unadjusted (no covariate adjustment)	1,546	0.160	0.0159	<0.001	0.770
Adjusted for baseline outcome measure only	1,546	0.148	0.0154	<0.001	0.714
Adjusted for all covariates	1,546	0.149	0.0196	<0.001	0.719
Alternative missing data treatment					
Adjusted for select covariates, with listwise deletion (with the same covariate specifications as the benchmark model)	1,346	0.151	0.0158	<0.001	0.728
Alternative estimation method					
Adjusted for select covariates, with imputed covariates, estimated with MLE	1,546	0.153	0.0117	<0.001	0.737
Alternative study sample					
Exclude two blocks that included schools that withdrew immediately after the random assignment.	1,457	0.157	0.0163	0.799	<0.001

Source: IMPAQ staff calculations based on student surveys.