Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

_____ _____

Rachel Catherine BeLieu Date

On Second Thought: An Evaluation of an Adolescent Pregnancy Prevention Program in Central Pennsylvania

By

Rachel Catherine BeLieu

Master of Public Health

Hubert Department of Global Health

_____

Dr. Roger Rochat

Committee Chair

On Second Thought: An Evaluation of an Adolescent Pregnancy Prevention Program in Central Pennsylvania

By

Rachel Catherine BeLieu

Bachelors of Science in Biological Sciences

University of Pittsburgh

2012

Thesis Committee Chair: Roger W. Rochat, MD

An abstract of

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in the Hubert Department of Global Health

2014

Abstract

On Second Thought: An Evaluation of an Adolescent Pregnancy Prevention Program in Central Pennsylvania

By Rachel Catherine BeLieu

Background: Formal sex education and adolescent pregnancy prevention have gone through several incarnations since first introduced in the 1970s. Over time, evaluation efforts have shown that comprehensive education programs, that include information on preventing STIs and pregnancy through the use of contraceptives, are more effective than programs that only promote abstinence. Furthermore, programs that teach social skills such as sex refusal, condom negotiation, and resistance of peer pressure are more effective than those only providing information.

Objective: To determine if significant changes in knowledge, attitudes, behaviors, intentions, and self-efficacy are seen in On Second Thought program participants between pre-test and post-test; to determine if changes from pre-test to post-test are associated with varying levels of intervention exposure and implementation fidelity; and to identify problem areas in program implementation and evaluation and recommend improvements.

Methods: On Second Thought is an adolescent pregnancy prevention program consisting of 7 90-minute sessions, covering topics of STIs, pregnancy, drugs and alcohol, relationships, self-esteem, and peer pressure. Program participants took pre-tests and post-tests assessing knowledge, attitudes, behaviors, self-efficacy, and intentions surrounding program topics. Facilitators also recorded program attendance, participation, and intervention fidelity. Post-test responses were compared to pre-test responses for participants to assess change, and change between pre-test and post-test was compared for those with differential intervention exposure.

Results: Greater intervention exposure was significantly associated with improvements in metrics relating to knowledge, attitudes, and behaviors. Students attending a greater number of successful sessions were more likely to show desirable change between pre-test and post-test on items dealing with STIs, condoms, and pregnancy. They were more likely to agree that early pregnancy would negatively affect their goals, and they were less likely to report using drugs or alcohol or having sex within the last 30 days.

Discussion: Despite poor data quality and a small sample size, this evaluation found significant effects of the intervention on a range of different metrics. Perhaps most importantly, intervention exposure was associated with changes in reported behaviors, a finding often not seen in evaluations of short-term, classroom-based pregnancy prevention programs such as this one.

On Second Thought: An Evaluation of an Adolescent Pregnancy Prevention Program in Central Pennsylvania

By

Rachel Catherine BeLieu

Bachelors of Science in Biological Sciences

University of Pittsburgh

2012

Thesis Committee Chair: Roger W. Rochat, MD

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in the Hubert Department of Global Health

2014

Acknowledgements

Without the support of an incredibly dedicated team of faculty and staff, this thesis would never have been written, and I would never have found myself where I am now, preparing for graduation.

I owe the greatest thanks to my advisor, Dr. Roger Rochat, and my ADAP, Angela Rozo, who never gave up on me even when they probably should have. Dr. Rochat has been the world's best possible mentor and source of inspiration, and the most steadfast pillar of strength and encouragement. Angela deserves all the laurels in academia for her tireless support, expert problem-solving, and passionate advocacy for my interests. I also want to thank Professors Stanley Foster and Mimi Kiser for their time, interest, and concern; and Professor Kate Winskell, without whose kind words I never would have made it past my first semester.

Many thanks are also due to the entire staff at the Family Health Council of Central Pennsylvania, who welcomed me with open arms and provided the inspiration for this thesis. I owe a great debt to Patricia Fonzi and Ayanna Cash-Clements, who brought me into their daily work and made me feel at home at the Council.

Finally, I must thank my family. Parenting terrifies me far more than this thesis ever did, but my parents and step-parents did an admirable job, especially considering the stubbornness of their charge. Great thanks to Mum, Brooks, Dad, and Tammy for supporting me through good times and bad, for cheering me on, and for raising me to be compassionate and curious. I have always known I am loved. Friends are a family we choose ourselves, and I owe many thanks to my chosen family for their care and love, for shared experiences and new ones. Last but not least, my thanks and my love to Sam Ziegler, who has brought joy to my life and stood by me patiently while I chase my dreams.

Contents

List of Figures

List of Tables

Chapter 1: Introduction

Section 1.1: Introduction and Rationale

Multiple social and individual factors affect exposure to risky behavior, the process of evaluating risks and making decisions, and the harms that can result from those decisions. On Second Thought hopes to act on the relationships between these factors and outcomes by building strong decision-making skills and by equipping adolescents to protect themselves from potential harms.

The program aims to prevent STI transmission, unintended pregnancy, and drug abuse among young adolescents in the Harrisburg metropolitan area in central Pennsylvania. The implementing organization, the Family Health Council of Central Pennsylvania (FHCCP) was interested in evaluating the impacts of this program after two years of implementation.

Section 1.2: Problem Statement

John Harris High School, in Harrisburg City School District, has a drop-out rate of 12%[1] and a 4-year graduation rate of 39%.[2] The state as a whole has a drop-out rate of 2% and a graduation rate of 83%. This puts the school in the bottom 5% of schools in the state. Unplanned pregnancy is a major reason why young women drop out of school. Nationwide, only 38% of women who gave birth before the age of 18 have a high school diploma. Of all female high-school dropouts, 30% cited pregnancy as the primary reason for their inability to complete school, and this rate increases to 36% of Hispanic and 38% of African-American girls.[3] The teen pregnancy rate in Dauphin County is 49 per 1,000 women 15-19; the birth rate is 37/1,000. This is significantly higher than the state as a whole; Pennsylvania has a pregnancy rate of 40/1,000 women 15-19 and a birth rate of 27/1,000.[4]

The county in general and Harrisburg in particular have much greater proportions of African-American residents than anywhere else in the central PA region. Harrisburg School District is 13% Hispanic and 55% African-American. Pennsylvania is 6% Hispanic and 11% African-American, in comparison. STD and

HIV rates in Pennsylvania are much higher in African-Americans as compared to Caucasians; the HIV rate is 12 times higher, Chlamydia is 14 times higher, and gonorrhea is 30 times higher. Dauphin County has a higher incidence of people living with HIV than the state as a whole (393/100,000 vs. 292) [5].

STD rates are always higher among young people; in Pennsylvania, the overall Chlamydia rate is 416 per 100,000 people; the rate is 2162/100,000 among 15-19 year-olds and 2279/100,000 in the 20-24 age group. A similar pattern is seen with gonorrhea. Furthermore, rates in Dauphin County are significantly higher than the state as a whole; the overall Chlamydia rate is 623/100,000, and the rate among 15-24 year-olds is more than 3000. Statewide, the gonorrhea rate is 108/100,000 overall; this is 241 in Dauphin County and exceeds 600 for the 15-24 group.

In Harrisburg City School District, 38% of all families with children, and 46% of individuals under 18, are below the poverty level [6]. This compares to 14% and 17% respectively, for Pennsylvania as a whole.

Section 1.3: Purpose Statement

The purpose of this project is to evaluate the effectiveness of the O2T program in positively changing knowledge, attitudes, intentions, and behaviors related to safer sex, pregnancy prevention, STIs and HIV, and drug use in adolescents.

Section 1.4: Objectives

Objective 1: Determine if significant changes are seen in program participants between pre-test and post-test.

Objective 2: Determine if changes from pre-test to post-test are associated with varying levels of intervention exposure and implementation fidelity.

Objective 3: Identify problem areas in program implementation and recommend improvements.

Section 1.5: Significance

Pennsylvania does not have a policy on sex education in public schools; periodic attempts to pass legislation to mandate comprehensive education have not succeeded.

The O2T program has not undergone any evaluation as of yet. Some information that could be used for monitoring and/or evaluation has been collected, but there has been little to no analysis of this data to assess the appropriateness of tools or the effectiveness of the program itself. FHCCP anticipates the availability of funding for rigorous M&E in fall 2014, and wishes to then begin the process of becoming certified as "evidence-based," pending the results of this study.

Process evaluation of this program will influence and improve future implementation. Fidelity monitoring, attendance and participation, and student feedback data will be used to identify problematic areas of implementation as well as effective strategies to maximize program impact. Fidelity monitoring data will help us identify which activities are being conducted as described in the curriculum, and which aren't. If certain activities are consistently being skipped or altered, we can look at which facilitators are doing the skipping or altering. If the problem seems widespread among different facilitators and sites, it may be necessary to modify that portion of the curriculum, using facilitator reports and student feedback as a guide.

Section 1.6: Definition of Terms

Sites, series, sessions, and activities:

A *site* or *host site* is a school, summer camp, community program, or other venue in which the intervention takes place, and from which participants are recruited. Some host sites are also the facilitators; for example, an FHCCP partner organization that hosts a youth program and uses the O2T

curriculum. Other sites are distinct from the facilitators; for example, a school at which a facilitator from a partner organization visits to conduct the intervention.

A *series* is one complete iteration of the intervention at a particular site. It consists of the sequence of seven sessions, and may occur over varying periods of time, from two weeks to four months.

A *session* is one of the seven units or lessons within the series. Each session focuses on a different topic area and is designed to be completed in about two hours. Typically, no more than one session is held each day within each series, but occasionally multiple sessions will be held in one day due to scheduling issues.

Activities make up the sessions. Each session consists of between four and seven separate activities, each of which has materials and facilitator scripts associated with them.

Chapter 2: Background and Literature Review

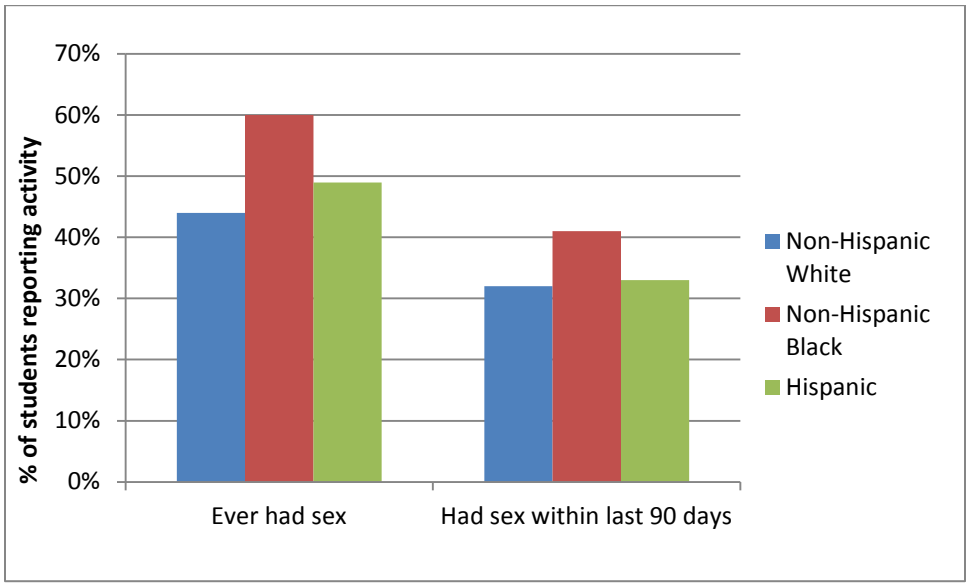
Section 2.1: Teen Sexual Activity

The 2011 Youth Risk Behavior Survey (YRBS) found that 47% of high school students reported having ever had sex. This varied significantly according to race and ethnicity, with 60% of non-Hispanic Black students reporting sexual activity compared to 49% of Hispanic students and 44% of non-Hispanic White students. By 12th grade, 66% of student had initiated sexual activity. These patterns were also seen for the proportion of students who were sexually active – i.e. they reported having sex within the last 90 days. Overall, 34% of students reported being sexually active, including 48% of 12th grade students [7].

According to the 2006-2010 cycle of the National Survey of Family Growth, the average age at first intercourse for adults age 15-44 was 17.1 for both men and women. In 2009 14% reported more than three sexual partners [8].

Sixty percent of students reported using a condom the last time they had sex, and 18% reported using birth control pills [7]. Eighteen percent of women ages 15-19 who are at risk for unintended pregnancy (those who reported having had sex within the three months prior to the interview) are not currently using any method of contraception. If we include women who are sexually experienced but did not have sex within the last three months, this proportion increases to 31% [9].

Figure 1: Sexual activity in high school students, 2011 YRBS



Section 2.2: Teen Pregnancy

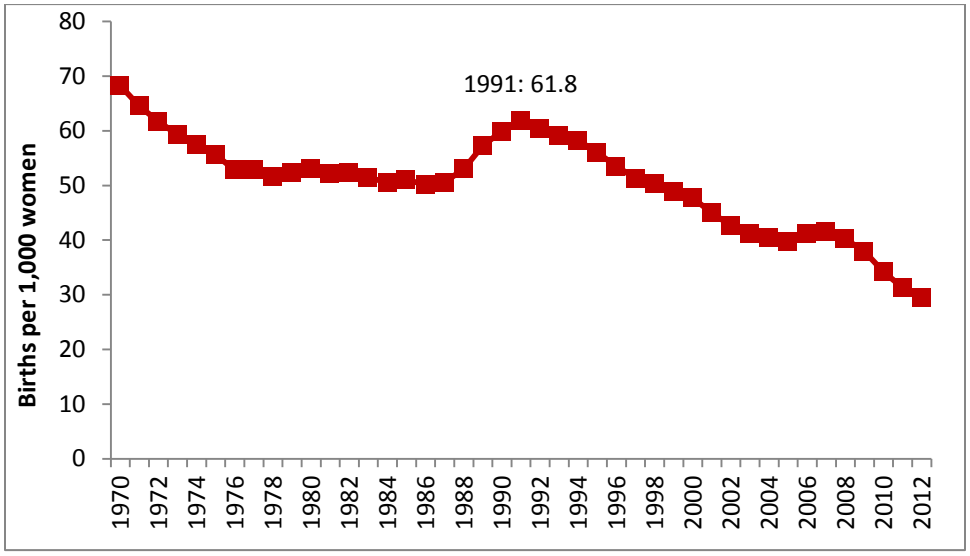
The teen birth rate has decreased significantly in the last 40 years. In 1970, the birth rate was 68.3 per 1,000 women ages 15-19. It decreased each year until 1979, and then stayed more or less stable – between 51 and 52 – until 1987. In the late 1980's the teen birth rate rose sharply, reaching a high of 61.8/1,000 in 1991. Since then, the birth rate has declined each year, aside from small increases in 2006 and 2007. Since 2007, the rate has dropped even more sharply, reaching an all-time low of 29.4/1,000 in 2012 [10]. Seventy-seven percent of births to women age 15-19 were unintended [11].

However, these rates differ greatly according to race. In 2012, the birth rate among non-Hispanic White women ages 15-19 was 20.5. The birth rate for Hispanic women of the same age was 46.3, while among non-Hispanic Black women it was 43.9 [10].

In Pennsylvania, the teen birth rate is lower than the national average, at 23.7/1,000. The highest teen birth rates occur in New Mexico (47.5), Oklahoma (47.3), Mississippi (46.1), Arkansas (45.7) and Texas

(44.4). Pennsylvania ranks 15th out of 50 states and the District of Columbia (with 1 being the lowest rate and 51 the highest) [10].

Figure 2: US Teen birth rate per 1,000 women age 15-19, 1970-2012

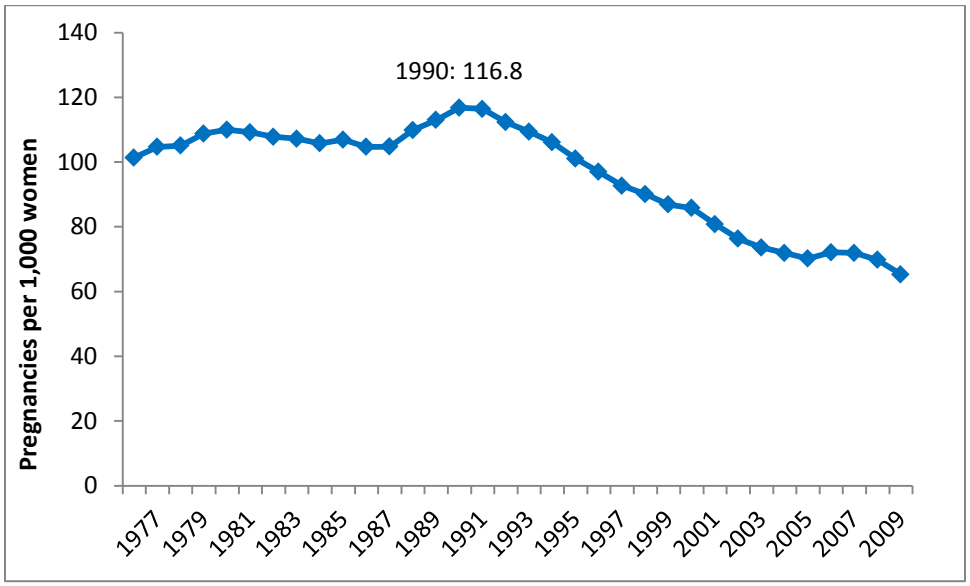


Teen pregnancy rates have followed a similar pattern over the years. Pregnancy rates stayed roughly the same from the late 1970s to the late 1980s, and peaked in 1990 at 116.8 per 1,000 women ages 15-19. Aside from a small increase in 2006, rates have dropped steadily since then, reaching an all-time low of 65.3/1,000 in 2009, the most recent year for which data is available. The rate of abortion has also steadily decreased, from 40.3/1,000 women age 15-19 in 1990 to 16.3 in 2009 [12].

Pregnancy rates also vary strongly by race. The 2009 pregnancy rate among non-Hispanic White women age 15-19 was 42.6 per 1,000 women. The pregnancy rate among their Hispanic counterparts was 100.1, and was 113.7 among non-Hispanic Black women of the same age [12].

The teen pregnancy rate in Pennsylvania in 2008 was 56 per 1,000 women age 15-19; lower than the national average, and ranking it at 14th out of 50 states and D.C. The states with the highest teen pregnancy rate include the District of Columbia (112), New Mexico (93), Mississippi (90), Texas (85), and Nevada (84) [13].

Figure 3: US Teen pregnancy rate per 1,000 women age 15-19, 1976-2009



Adolescent child-bearing carries risks not only for young mothers, but for their babies. Children born to a teen parent are more likely to experience delays in cognitive development, educational failure, and behavioral problems; and are more likely to become teen parents themselves [14].

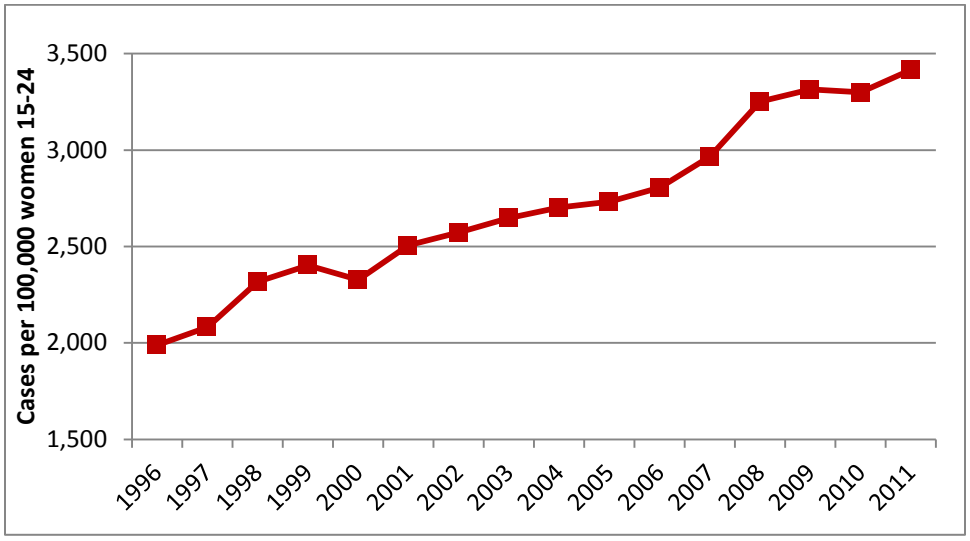
In 2010, the public costs associated with teen childbearing were estimated at $9.4 billion; however this represents a decrease of $1.5 billion in estimated costs from 2008. The declining teen birth rate is likely responsible. In fact, the decrease in the adolescent birth rate between 1991 and 2010 is associated with savings to taxpayers of $12 billion in 2010 alone [15].

Nationwide, only 38% of women who gave birth before the age of 18 have a high school diploma. Of all female high-school dropouts, 30% cited pregnancy as the primary reason for their inability to complete school, and this rate increases to 36% of Hispanic and 38% of African-American girls [3]. If every member of the high school class of 2009 had graduated, the US could have avoided $320 billion in lost economic activity and $174 billion in excess healthcare costs [15].

Section 2.3: STIs among Adolescents

DiClemente and colleagues comment that “the risk of acquiring an STD is one of the most substantial and immediate threats to the health and well-being of adolescents” [16]. Chlamydia infection has skyrocketed in recent years, increasing from 6.5 cases per 100,000 in 1984 to 458 cases per 100,000 in 2011. Chlamydia rates are particularly high among young women; in 2011, women ages 15-24 saw 3416 cases per 100,000 [17]. Each year, 20 million new STI infections are diagnosed in the United States, generating annual medical costs of $16 billion, according to a 2013 estimate. There were more than 110 million new and existing cases of STIs as of 2008, including nearly 80 million cases of HPV and more than 1.5 million cases of Chlamydia [18].

Figure 4: Chlamydia incidence, women age 15-24, 1996-2011



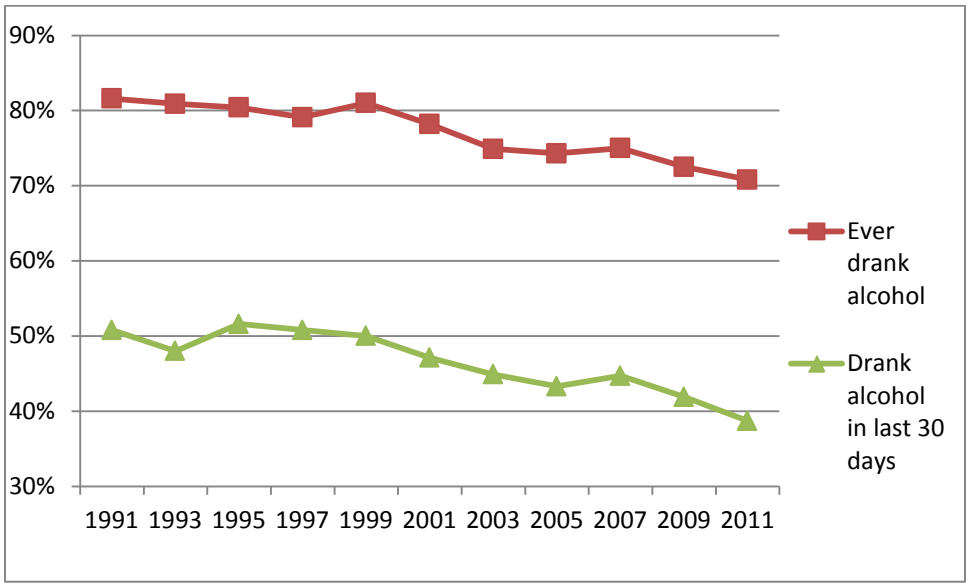
In 2012, Chlamydia infections among those under 25 represented 70% of all cases reported that year.[19] A quarter of young women ages 14-19 in the US are infected with at least one STI – representing 3.2 million adolescents [20].

Section 2.4: Drug Use among Adolescents

In 2011, 71% of high-school students reported ever drinking alcohol in their lives; 39% reported doing so within the last 30 days. Furthermore, 22% reported binge drinking (five or more drinks of alcohol within

a couple hours) within the last 30 days. In the time that the Youth Risk Behavior Survey has been conducted, alcohol use has decreased markedly. In 1991, 82% of high school students reported ever drinking alcohol while 51% reported use within the last 30 days and 31% reported binge drinking [21].

Figure 5: Alcohol Use among US High School Students, 1991-2011



On the other hand, marijuana use has increased from a low in 1991, though much diminished from its peak in 1999. In 2011, 40% of high school students reported ever using marijuana, and 23% reported doing so within the last 30 days. Use increased significantly from 2009 to 2011. This compares to 31% reporting ever-use and 15% reporting last-30-day use in 1991; in 1999, 47% reported ever-use while 27% reported using within the last 30 days [21].

In 2011, 45% of students reported ever having tried smoking cigarettes; a marked decreased from a high of 71% in 1995. The same pattern is seen in cigarette smoking within the last 30 days: in 2011 18% reported smoking on at least one day during the last month, while 6% reported smoking on at least 20 days. These figures have decreased from their peak in 1997, when they stood at 36% and 17%, respectively [21].

Section 2.5: Interventions

2.5.1: History and Policy

The proportion of students receiving some form of sex education increased from 47% of 17-year-olds in 1979 to 90% in 1994,[22] and to over 97% in 2008 [23].

Approaches to sex education interventions have evolved over time, with programs falling into one of four distinct "generations." The first generation of interventions, implemented in response to rising rates of STIs and teen pregnancy, relied heavily on "scare tactics" emphasizing the dangers of "venereal disease." The second generation incorporated more information on contraception and safe sex, and the third generation saw a return to an abstinence-only approach as several federal funding programs in the 1990s and early 2000s provided grants exclusively for this type of education. The fourth generation of sex education programs involved a greater emphasis on building social and behavioral skills to enable safer sex practices, such as resisting pressure, refusing sex, and negotiating condom use [24].

In 1981 the Adolescent Family Life Act was signed by President Reagan. This law marked the first time that the federal government invested in local pregnancy prevention programs. The goal of the program was to promote "chastity and self-discipline" as the most important strategy to prevent teen pregnancy [25].

In 1996, another law dealing with sex education was passed as part of a larger welfare reform act. This law invested $440 million in abstinence-only sex education programs over five years. This law created the strict standards that govern abstinence-only education – curricula must "condemn *all* sex acts outside of marriage – for people of any age" and may not discuss contraception in any positive way. Contraception may only be discussed to highlight its risks and failure rates [25]. Studies show that over the 1990's, teachers and schools became more likely to emphasize abstinence and less likely to include discussion of other methods of contraception. From 1988 to 1999, the proportion of sex education

teachers who presented abstinence as the only method of preventing STIs and pregnancy increased from 2% to 23%. In the same vein, the proportion of teachers who considered abstinence their most important message rose from 25% in 1988 to 43% in 1999 [26]. In 2000, another $50 million over two years was approved by Congress through the maternal and child health block grant [25].

The years of the George W. Bush administration saw an enormous increase in the amount of federal funding directed towards abstinence-only education – that is, curriculums that must focus solely on sexual abstinence until marriage as the only effective way to prevent pregnancy and STIs. Programs may only discuss other forms of contraception to describe their rates of failure. From 1995 to 2002, the proportion of teenagers who received instruction on contraception decreased from 84% to 65% in females and from 65% to 59% in males. The proportion who only received abstinence-only education increased from 8% to 21% in females and 9% to 24% in males. Furthermore, 38% of males and 30% of females in 2002 reported receiving no formal sex education at all [27]. Between 2001 and 2004, federal spending on abstinence-only education increased from $80 million to $137 million. By 2007 this figure had risen to $204 million [22]. In 2005, one-third of US school districts required sex education to be "abstinence-only" according to the definition above [26].

During the Obama administration, the national focus shifted towards comprehensive sex education. The first budget proposed by the administration, for fiscal year 2010, replaced much funding for abstinence-only education with funding for comprehensive sex education [28]. The National Health Education Standards are part of the Healthy People 2020 Initiative. One of the goals of Healthy People 2020 is to increase the provision of high-quality health education in schools. Among the relevant goals in these standards are:

- Standard 1: Students will comprehend concepts related to health promotion and disease prevention to enhance health.

- Standard 4: Students will demonstrate the ability to use interpersonal communication skills to enhance health and avoid or reduce health risks.
- Standard 5: Students will demonstrate the ability to use decision-making skills to enhance health.
- Standard 6: Students will demonstrate the ability to use goal-setting skills to enhance health [8].

As of 2014, twenty states and the District of Columbia mandate both general sex education and HIV education. Another two mandate sex education only, and another 13 mandate HIV education only [29]. Only 18 states and D.C. require that sex education, if taught, provide information on contraceptives; while a total of 37 states require that education provide information on abstinence. Twelve of these states require just that abstinence be covered, and 25 require that abstinence be stressed. Nineteen states require that instruction include "the importance of engaging in sexual activity only within marriage" [29].

2.5.2: Current State

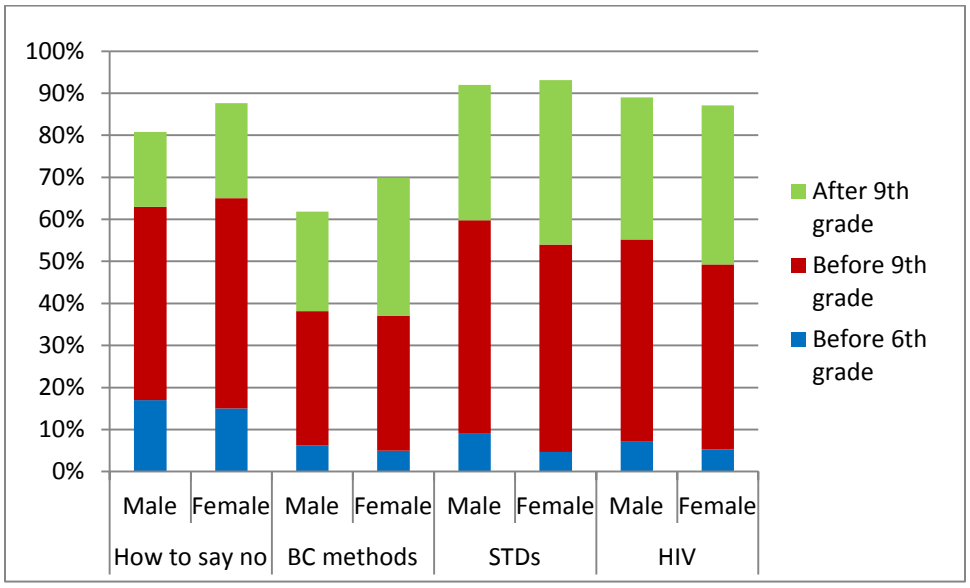
While 96% of female and 97% of male teenagers reported receiving formal sex education before the age of 18 [23], a study using data from the 2006-2008 cycle of the National Survey of Family Growth examined in particular what sex education teenagers received *before* sexual debut (first vaginal sex). The study found that about two-thirds of females and a little over half of males received comprehensive education, while about 20% of each received abstinence-only education [27].

Table 1: Receipt of Sex Education

	Female	**Male**
Comprehensive	66%	55%
Abstinence-only	18%	21%
None	16%	24%

This is shown by other results from the NSFG, which demonstrate that many young people do not receive sex education in important areas until later in adolescence. Less than 40% of male and female adolescents had received education in methods of birth control before entering the 9th grade. The same was true for education on STDs (less than 60%) and HIV (around 50%) [23]. Considering that 14% of female and 18% of male adolescents reported having sex before the age of 15, waiting until high school to provide sex education is insufficient [30].

Figure 6: Proportion of adolescents receiving sex education in different topics



2.5.3: Theories and Models

Social norms theory and the theory of reasoned action states that behavior is influenced by individual perceptions of peer or group norms. For example, adolescents are more likely to begin having sex if they perceive that their friends or classmates are doing so [16]. Creating a space within an intervention in which alternative decisions can be discussed and considered, particularly while in the company of one's peers, may allow for shifts in perception of social and behavioral norms.

DiClemente and colleagues propose an ecological approach to STI and HIV interventions; interventions at the "Mesosystem" level, which involve the immediate settings and relationships of adolescents.

Interventions using this strategy must go beyond simple transmittal of knowledge and enable adolescents to safely and successfully navigate the challenges presented by the environment. Particularly important is inculcation of skills in refusing sex or negotiating condom use. The authors emphasize that traditional knowledge-centered approaches focus too closely on the level of the individual, and neglect the larger social and cultural forces that shape ability to control one's sexual experiences. For this reason, young women are especially ill-served by interventions that fail to teach behavioral skills for coping with the risks presented by this environment [16]. Interestingly, other studies have found that changes in behavioral and health outcomes such as experiencing or causing pregnancy, contracting STDs, having unsafe sex, early sexual debut, and multiple sexual partners are seen in young men but not in young women. It may be that interventions that focus only on knowledge or only on individual decision-making do provide important skills and resources for young men to avoid risk, but this content is not sufficient to enable women to negotiate safer sexual experiences.

Gavin et al. describe an approach to sex education that combines "Positive Youth Development" (PYD) with the knowledge and skill-based content of traditional sex education. This approach "target[s] a different, but complementary, set of mediating variables" that determine pregnancy and STI acquisition. PYD programs include at least one of twelve distinctive program goals, than include cognitive, social, behavioral, and emotional competence, self-determination, self-efficacy, and prosocial norms. These programs are also empower participants to develop their goals for the future and practice self-determination. The most effective PYD interventions are of relatively long duration (at least one entire school year), to better promote the formation of positive relationships and the gradual changes the program hopes to create [20]. O2T addresses many of these goals with its emphasis on thoughtful decision-making; adolescents are encouraged to consider risks and benefits of different actions on an individual basis, making their own determinations based on their identified values and future goals.

PYD programs began with a strong focus on a strengths-based approach, bucking the usual trend of programs that see "youth as troublesome and in need of fixing." There was initial reluctance to address prevention or risk reduction because these topics are usually associated with a worldview that "characterize[s] adolescents as having problems." However, recently there has been more integration of the positive approach of PYD programs with components that focus on reduction of risky behaviors. Implementers found that these two approaches were not antithetical but could be united to provide a more thorough and effective prevention strategy [31].

Section 2.6: Evaluation of Interventions

2.6.1: Studies Using National Data

There is a dearth of national data on sex education, particularly when it is offered during adolescence. Estimating the impact of sex education programs on health outcomes is difficult because so many other factors affect adolescent sexual behavior and risk-taking [22].

Analysis of some data, such as the National Longitudinal Study of Adolescent Health, may lead researchers to "the naïve conclusion that's sex education causes adverse health outcomes," because receipt of sex education and sexual risk-taking often rise in concert. However, after controlling for other variables, there is no association between receipt of comprehensive sex education and earlier sexual debut, more sexual partners, or greater frequency of sex [22].

The study using data from the 2006-2008 National Survey of Family Growth found a strong effect of sex education on sexual debut, with little difference between comprehensive and abstinence-only; around 87% of those with no education reported having had sex by age 20, compared to about 77% of those with some sex education [27].

Among females, receipt of comprehensive sex education was significantly associated with improvements in five different sexual health outcomes, while abstinence-only education had no significant effects. Furthermore, for two factors – whether first sex was with age-discrepant partner and whether condoms were used at first sex – the difference between comprehensive and abstinence-only education was significant. For males, comprehensive education was significantly associated with improvements in eight health outcomes, and abstinence-only education in five. In no outcomes was there a significant difference between comprehensive sex education and abstinence-only. In addition, the strength of association between receipt of some education (whether comprehensive or abstinence-only) and positive outcomes was stronger in males than in females for almost all factors [27].

The results of the National Survey of Family Growth study show that receipt of comprehensive sex education over abstinence-only education is especially important for young women. Interestingly, the study showed a very strong and significant negative association between comprehensive sex education (though not abstinence-only) and first sex being unwanted – though only in females. The differences between abstinence-only and comprehensive for first sex being with an age-discrepant partner were significant for women. In fact, abstinence-only education made an age-discrepant first sex partner slightly *more* likely, though this association was not significant [27]. Sex education may be a particularly gendered issue. It is telling that abstinence-only education was sometimes effective for young men but almost never for young women. The skills and knowledge taught in abstinence-only are clearly in adequate for young women navigating sexual maturity.

2.6.2: Meta-Analyses

A 2005 review of 16 Randomized Control Trials showed that most studies had short follow-up periods, and few assessed end-point behavioral or health outcomes, such as occurrence of pregnancy. Most of the evaluation indicators, rather, were “surrogate markers” for pregnancy. The researchers argue that

intention measures, such as intention to use contraception or to abstain from sex in the future, are inadequate outcome indicators for evaluating sex education programs. However, they (somewhat controversially) included contraceptive knowledge in their comparison of intervention outcomes, in addition to self-reported sexual debut, number of partners, frequency of sex, and use of contraception [26].

Four out of five studies found a significant improvement in contraceptive knowledge compared to controls, and another study also found significant improvement in comparison to an abstinence-only program. Similarly, of the 10 studies evaluating contraceptive use, five saw significant increases in use or decreases in unprotected sex compared to controls, and another showed the same significant increase when compared to the abstinence-only program [26].

Gavin and colleagues' evaluation of Positive Youth Development (PYD) interventions focused solely on behavioral and health outcomes, such as engagement in intercourse, use of birth control, frequency of unprotected sex, number of sexual partners, history of pregnancy, and diagnosis with an STI.[20] They did not include knowledge or attitude outcomes, which the majority of test items in the O2T program represent.

Evaluation of 30 PYD programs found that 15 of these improved at least one sexual of reproductive health outcome; 6 increased use of birth control, 3 decreased frequency of sex, and 7 delayed sexual initiation. In addition, 9 programs reduced substance use. Researchers found that programs that incorporated more characteristics of PYD interventions were more effective at achieving change in behavioral or health outcomes. The authors did note that replication of the most effective programs has been difficult, and pointed out that implementation fidelity is often unmeasured in such evaluations, and can have significant impacts on the failure or success of programs. A limitation of this study was the lack of information on program dosage levels [20].

2.6.3: Evaluation of Individual Interventions

The Adolescent Decision-Making for the Positive Youth Development Collaborative (ADM-PYDC) is an 18-session curriculum that aims to prevent substance abuse through teaching prevention skills and conducting health education. Seven sessions are spent on effective decision making, two on information on the risks of alcohol and drug use, and four on applying this decision-making strategy to one's life. This included "identifying positive personal attributes" [31].

The outcome variables of the study included drug use attitudes, perceptions of risks of drug use, and self-reported recent use of substances. Attitudes and risk perceptions were measured by four Likert-scale items each, while recent use of measured by a series of questions asking if certain substances had been used in the last 30 days. Participants who attended at least 50% of sessions were compared to those who attended less than that; however there were no significant differences between these groups and so an "intention to treat" analysis was performed instead [31].

The intervention group saw a significant increase in perceptions of harms of drugs from pre-test to post-test compared to the control group. Alcohol use decreased from pre-test to follow-up for the intervention group and not the control group. Both groups saw increases in marijuana and other drug use from pre-test to follow-up, but the intervention group showed significantly smaller increases than the control group in both cases. This can still be considered a positive outcome, as research shows that slowing the rate of increase in substance use is protective against later increases in use [31].

The evaluation of the Wise Guys curriculum used identical survey instruments for the pre-test, post-test, and six-month follow-up tests. Among other things, each test included 13 items assessing general knowledge of sexuality and reproduction, 8 items assessing STI transmission knowledge, and 14 items assessing desirable attitudes toward sexuality and sexual behavior. Both True/False and Likert Scale variables were recoded as dichotomous variables. Participants saw increases in correct responses to all

13 general knowledge items (compared to 6 items in controls), 6 of 8 STD knowledge items (compared to 2 items in controls) and all 14 desirable attitudes (compared to 5 items in controls). Participants typically saw slight "decay" in knowledge and attitude items from post-test to follow-up; scores on both post-test and follow-up were significantly higher than controls, as were increases from pre-test to post-test [32].

Participants also saw statistically significant increases in composite variables for all three sections compared to controls. The instrument also measured sexual behavior with three questions assessing sexual debut, use of contraception at last sex, and frequency of contraception use. Among those who initiated sexual activity between pre-test and follow-up, participants were significantly more likely than controls to report using contraception "every time" they had sex, and more likely to report use of condoms at last sex, at follow-up. Sexually active participants were also more likely than sexually active controls to begin using condoms between post-test and follow-up. No other differences in behavior between participants and controls were statistically significant. Regression analysis found that participation in the program was the most significant predictor of general knowledge and STD knowledge at post-test and follow-up; the same was true for use of contraception "every time" [32].

An evaluation of the Prime Time clinic-based pregnancy prevention program for adolescent girls found that participation in the intervention was associated with significantly more consistent use of condoms and hormonal contraception with the last sexual partner at the twelve-month interim evaluation. While the evaluation also assessed psychosocial outcomes, such as attitudes, beliefs, and perceived norms surrounding sexuality and contraception use, little change was seen in these measures after twelve months. Participants did show significant improvement in stress-management skills and school and family connectedness, both of which are associated with increased contraception use [14].

Many evaluations of fourth-generation programs have found significant impacts on attitudes and knowledge, but not on behavior. Programs such as Reducing the Risk and Be Proud! Be Responsible! are among the few showing significant changes in behavioral outcomes. A quasi-experimental study by Zimmerman and colleagues tested an approach to these interventions that specifically targets participants with sensation-seeking and impulsive personality characteristics, which in health behavior models are connected to risk-taking behavior. Zimmerman and colleagues hypothesized that interventions based on a model of purely rational health decision-making will fail to produce expected changes in behavior. However, the study in Cleveland, OH and Louisville, KY public schools failed to support this hypothesis. Both the standard Reducing the Risk curriculum and the enhanced curriculum altered to appeal to sensation-seeking and impulsive personality traits saw significant reductions in sexual initiation compared to a control group. Both curriculums also saw significantly increased knowledge, between baseline and end line, though these differences were no longer significant at follow-up. However, there were no significant differences between the standard and enhanced curriculums on any outcome measures [24].

The researchers in the Reducing the Risk study found that several mediating variables were significantly associated with initiation of sexual activity, including attitudes towards waiting to have sex and intentions to have sex [24]. Many meta-analyses have discredited the use of mediating or psychosocial variables as evaluation outcomes, claiming that only behaviors such as frequency of sexual activity or use of contraception are useful metrics for assessing program impacts on adolescent sexual health. The results of this study show that some of these mediating variables are significantly associated with behavioral outcomes, and can be useful measures of program impact.

Reviews have shown Reducing the Risk and Be Proud! Be Responsible! to be consistently effective in increasing condom use, reducing unprotected sex, increasing contraceptive use (RtR only) and reducing the number of sexual partners (BPBR only) [8].

Research has shown that effective interventions are based on proven theories of health education. The most effective programs addressed issues such as knowledge, perceived risks and norms, values, attitudes, and self-efficacy [8]. Short-term programs (such as O2T) unfortunately have been shown to have little to no impact on sexual behavior, especially when compared to intensive, long-term curricula. Furthermore, the most effective interventions expanded their focus beyond information provision to cultivate negotiation and refusal skills in teens [22]. Most studies have found no association between abstinence-only programs and positive behavior change or reduction of risk factors. On the other hand, a seminal 2007 review found that two-thirds of comprehensive education programs studied had positive effects on sexual risk behavior [27] [33].

DiClemente and colleagues identified antecedents of STI and HIV infection, including perceived risk of infection, self-efficacy, self-esteem, condom attitudes, condom use expectancies, and peer norms surrounding sexual behavior. They found that interventions that address these antecedents and focus on decision-making skills and social capacity were the most effective in reducing risk of STI or HIV infection [16].

Chapter 3: Methods

Section 3.1: Context

On Second Thought (O2T) was created and implemented by the Family Health Council of Central PA (FHCCP) in 2011. FHCCP is a 501(c)(3) non-profit organization with an annual budget of approximately $12 million. Revenue mainly comes from federal and state funding sources, with the vast majority through the Title X Family Planning Program, Ryan White HIV/AIDS Program, and WIC Supplemental Nutrition Program. Founded in 1973, FHCCP provides reproductive, sexual, maternal, and child health services in a 24-country area of central PA. The organization is one of four entities in the Alliance of Pennsylvania Councils, which together provide services for the entire state. FHCCP accomplishes most of its work through a network of partner organizations that individually provide services, but also provides direct services through several Tapestry of Health sites throughout the region.

The project was initially conducted in three school districts in Dauphin County. Dauphin County is positioned at the junction between several cultural and geographic regions of Pennsylvania. Most of the population lives in the central and southern portions of the county; the northern areas are thinly populated. Philadelphia is less than 100 miles to the southeast; nearer at hand lie the smaller cities of Lancaster, York, and Lebanon, which have large Hispanic and transient worker populations. North of the county, Pennsylvania's Northern Tier begins, a very rural, sparsely populated, generally poor, and overwhelmingly White region. Directly west lies affluent, suburban Cumberland County.

The county includes the state's capitol city, Harrisburg, and its eastern suburbs, including Hershey. Harrisburg is situated directly on the eastern shore of the Susquehanna River, which comprises the county's western border. The city of Harrisburg is only the 9th largest in the state, with a population of about 50,000; however about 550,000 people live in the metropolitan area [34]. Dauphin County differs

from its neighbors in having the only significant African-American population, at 17%, which is largely concentrated in Harrisburg [35].

In 2010, Harrisburg was rated by Forbes the second-best place to raise a family in the US, [36] and by Newsweek as among the most “recession-proof” cities [37]. However, city government has been beset by financial difficulties. The city’s total debt is estimated at $1.5 billion [38]. Harrisburg was the first municipality ever to be charged with securities fraud[39], and was blocked from filing for bankruptcy in 2011 [40].

Harrisburg City School District has its own difficulties, with $437 million in long-term debt, and consistently very low test scores. In 2013 it was ranked 493 out of 500 public schools in PA based on standardized test scores [41], and only one school of 11 in the district achieved adequate yearly progress [42]. Governance of the district has bounced between an independent school board and the mayor over the last 10 years [43].

Central Pennsylvania is politically and socially conservative for the most part. Urban areas, as usual, are somewhat more liberal. The PA governor, Tom Corbett, is a Republican, and Republicans hold majorities in both chambers of the Pennsylvania General Assembly. The mayor, city controller, and state legislative representatives from Harrisburg are all Democrats. The implementation of this program has already been affected by these factors; leaders of a summer enrichment program refused to allow discussion of more controversial topics, and one direct service site was unable to find anyone who would agree to host the sessions.

Section 3.2: Intervention

O2T’s theoretical basis draws from the Theory of Reasoned Action and the Health Belief Model. It is also based on the concept of “self-exploratory learning” and Howard Gardner’s Theory of Multiple Intelligences.

The Theory of Reasoned Action asserts that behavioral intentions depend upon both an individual's attitude towards a particular behavior and perceived subjective norms derived from the social context. O2T provides the opportunity for participants to develop their own attitudes about behaviors in a supportive environment. The Health Belief Model posits that health decision-making is predicated on four factors: perceived susceptibility, perceived severity, perceived benefits, and perceived barriers. O2T aims to influence all four factors by providing accurate information on the risks and dangers of STIs and drug use, for example, and fostering attitudes that help students see benefits and overcome barriers. The curriculum uses the Theory of Multiple Intelligences by presenting content in a variety of formats, engaging students whose learning styles and strengths differ.

On Second Thought (O2T) is designed for adolescents between the ages of 9 and 18; students in middle or high school. It was developed over a period of several months by FHCCP in partnership with an external consultant. The curriculum consists of a series of seven sessions. Each session covers a different theme or topic, and involves three to six different activities. The curriculum is intended for groups; many of the activities work best with at least four or five participants. Sessions with up to 30 participants have been conducted successfully; however additional educators/facilitators are required. The seven sessions are as follows:

- Session 1: Introduction and Self-Esteem
- Session 2: Puberty and Teen Pregnancy Prevention
- Session 3: Preventing Pregnancy and Promoting Health Relationships
- Session 4: Relationships and Drugs and Alcohol
- Session 5: STDs and HIV
- Session 6: Role-Play and Project Prep
- Session 7: Special Projects and Closure

The O2T curriculum guide identifies the objectives of the program: “Using a self-exploratory approach, youth will develop attitudes and skills to assist them in making healthy and informed decisions regarding sexual behaviors. As a result of participating in *O2T*, students will have:

1. Increased confidence in their ability to make informed decisions regarding sexual behavior.
2. Articulated their knowledge of behaviors associated with healthy sexual relationships and behavior.
3. Evaluated the risks of adolescent sexual behaviors and drug/alcohol use.
4. Demonstrated the ability to avoid negative or unwanted sexual behaviors.
5. Identified a support system to help them apply lessons learned from *O2T* in their lives”

The topics covered include:

- Male and female reproductive anatomy
- Puberty
- Biology of intercourse, conception, and disease transmission
- Mode of action, effectiveness, and use of different methods of birth control and contraception
- Proper condom use
- Risks of disease transmission and pregnancy
- HIV/AIDS, different STDS and their symptoms
- Alcohol and drugs and their risks
- Peer pressure and refusal skills
- Self-esteem and self-worth
- Healthy romantic and intimate relationships
- Emotional readiness for sex

- Long-range planning and goal-setting for the future
- Thoughtful decision-making

The evaluation of this program was undertaken by FHCCP within the unique political climate following passage of the Affordable Care Act. Both the ACA and the 2009 Recovery and Reinvestment Act provided billions of dollars in new funding for health. The Department of Health and Human Services placed an increasing emphasis on evidence-based programs, and there were multiple grant opportunities for research and demonstration projects. In September 2010, OAH awarded $15 million from the newly-created Teen Pregnancy Prevention Program to implement and evaluate new program models [44]. The Administration on Children and Families awarded $44 million in FY 2010 through the State Personal Responsibility Education Grant Awards, created by ACA [45], and the Office of Minority Health was awarding grants for demonstration projects through the Youth Empowerment Program [46].

FHCCP was eager to take advantage of this funding environment. In addition, FHCCP was already conducting an adolescent education program, called Teen Game Plan, but funding for this program was coming to an end.

Section 3.3: Study Design

This study is a quasi-experimental plausibility evaluation, using a reflexive control group along with a dose-response analysis. This program was not implemented with eventual evaluation in mind; no comparison or control groups were established.

For each student, there is a record of their attendance at the different sessions in the series, as well as whether activities at each session were conducted (at all or according to the curriculum), and whether students in that class participated and were interested in the activities of each session. With this information, we can compare students within the overall group of participants who have received

varying levels of exposure to the intervention. Students take pre- and post-tests assessing knowledge, attitudes, intentions, and behaviors surrounding safe sex, drug use, and intimate relationships. The intervention intends to create positive change in these measures. Both the end scores (at post-test) and the change between pre- and post-test scores can be compared for students with high levels of exposure and students with low levels of exposure.

3.3.1: Participants

This research considers 213 students enrolled at three different sites and 11 different series. The number of students enrolled at each site ranged from 5 (at the Summer 2012 New Steelton site) to 38 (at Ashler Manor). The average number of enrolled students per site was 19.4. The students ranged in age from 12.0 to 20.9 years. The students were mostly non-White, with 70 of the 102 students with race information recorded identifying themselves as Black or Hispanic alone, and only 6 identifying themselves as White alone. More details on the numbers and characteristics of participants are found in the Results chapter.

3.3.2: Measures

Process indicators measure whether the central intervention – the 7-session educational program – was conducted as intended. Data collected include sessions held, activities conducted within sessions and fidelity to curriculum for those activities, attendance, participation, and student satisfaction.

Facilitators currently collect most of this information; at each session they take attendance and also complete a fidelity monitoring log, recording whether each activity in that session was conducted as described, with changes, or not at all. They include notes on why this was the case (e.g. not enough time, insufficient attendance, forbidden by hosting site). For the session as a whole, they record whether the time was sufficient, and whether students were interested and participated in the activities, as well as additional notes on student behavior or other problems.

- Number of sessions held
- Date of session and session length
- Number of students in attendance at each session
- Number of activities in each session conducted as described in curriculum or with changes
- Student participation and interest for each session
- Whether session felt rushed

Pathfinder International uses similar measures, including number of courses held, number of youth attending or completing courses, number of youth participating in a training program, number of training sessions held, and number of favorable reports on sessions by participants [47]. Chinman et al describe how monitoring indicators can measure adherence, dosage, quality, and response by observing adherence to curriculum or model, teaching style and skill, and student participation; and by recording the number of modules each participating youth receives as part of the intervention [48].

Outcome indicators are the responses of program participants to questions on the post-test, and the change in their responses between the pre-test and the post-test. The tests consist of twenty questions addressing knowledge, attitudes, intentions, self-efficacy, and behaviors surrounding drug use, sexual activity, intimate relationships, pregnancy and STI/HIV prevention, and peer pressure.

- Proportion of students responding “correctly” to questions on post-test and pre-test
- Average scores on questions on post-test and pre-test
- Average numeric change in score from post-test to pre-test
- Proportion of students moving in a “positive” direction from pre-test to post-test

These indicators are similar to those used in the evaluation of adolescent reproductive health programs elsewhere. A large set of indicators developed by Pathfinder International includes measures that fall into the same categories. Other studies conducted in the United States ask about frequency of unprotected sex[20] as well as alcohol and drug use in the last month or last 90 days [49]. Jemmott et al assessed beliefs about the use of condoms to prevent pregnancy and/or disease, self-assessed ability to negotiate condom use and use a condom effectively, intention to use condoms, and correct knowledge about condoms. Other measures included knowledge of the transmission and consequences of STIs/HIV, and perception of abstinence's effectiveness in preventing pregnancy/disease [50].

Boyce et al, in a Canadian study, asked whether participants planned to use a condom, would ask a partner about condom use before sex, would consent to sex with a partner who refused to use a condom, and if they felt confident in their ability to use a condom [51]. Howard's study of sex education for young men tested knowledge by asking if "a girl can get pregnant the first time she has sexual intercourse" and if "some [STDs] have no symptoms." Attitudes and intentions were measured by asking whether respondents considered it appropriate for young teens to have sex [52].

3.3.3: Instruments

Intervention exposure was measured by the facilitator. At the first session, the facilitator recorded the name and date of birth of each participant. At each session, attendance was taken and the facilitator recorded whether each participant was present.

At each session, the facilitator also completed a fidelity monitoring form. On this form they reported the date(s) of the session and its length, the number of students present, whether students were interested in the session and whether they participated, and whether the session felt rushed. In addition, for each activity within the session, the facilitator recorded whether the activity was conducted as described in

the curriculum, with changes from the curriculum, or not at all. Copies of the fidelity monitoring forms for all seven sessions, in addition to facilitator instructions, can be found in Appendix 1.

The clear weakness of this method is that it depends upon complete and honest reporting from facilitators. Some facilitators may be hesitant to report that they did not stick to the curriculum, or they may simply recall the session incorrectly. On the other hand, with adequate training, the possibility of facilitator bias in recording session details could be reduced. So far, facilitators have frequently reported "negative" results, suggesting that this method may be suitable.

Knowledge, behavior, attitudes, intentions, and self-efficacy were assessed using a 20-item survey. This survey was used for both the pre-test and the post-test, with no changes. Seven items had Yes/No response options, and thirteen had a Likert-type scale with five possible responses, ranging from "Strongly disagree" to "Strongly agree." The middle response was "Don't know," with no option for "neutral" or "neither."

Table 2: Pre/Post-test questions, response values, and desired responses

Question	Response Values	Desired Response
1. Can a girl get pregnant the first time she has sex?	Yes/No/Don't Know	Yes
2. Can you get STI's from having oral sex?	Yes/No/Don't Know	Yes
3. Can you get HIV from kissing someone?	Yes/No/Don't Know	No
4. Can a person have an STI and not know it?	Yes/No/Don't Know	Yes
5. Can you say "no" to sex if your love that person?	Yes/No/Don't Know	Yes
6. Can some drugs kill you, even if you use them one time?	Yes/No/Don't Know	Yes
7. Can wearing a condom keep you from getting an STI or HIV?	Yes/No/Don't Know	Yes
8. In the past month I have had sex (vaginal, anal, or oral.)	Strongly Agree/Agree/Don't Know/Disagree/Strongly Disagree	Disagree
9. Getting pregnant or getting someone pregnant at my age would make it more difficult to achieve my goals.	Strongly Agree/Agree/Don't Know/Disagree/Strongly Disagree	Agree
10. Having sex would be a good way to keep a boyfriend or girlfriend.	Strongly Agree/Agree/Don't Know/Disagree/Strongly Disagree	Disagree
11. I am too young to do drugs or drink alcohol.	Strongly Agree/Agree/Don't Know/Disagree/Strongly Disagree	Agree
12. I know how to use a condom.	Strongly Agree/Agree/Don't Know/Disagree/Strongly Disagree	Agree
13. I feel that I have a number of good qualities.	Strongly Agree/Agree/Don't Know/Disagree/Strongly Disagree	Agree
14. I would insist that my partner or I use a condom every time we had sex.	Strongly Agree/Agree/Don't Know/Disagree/Strongly Disagree	Agree
15. If a person has unprotected sex one time, s/he can get an STI or HIV.	Strongly Agree/Agree/Don't Know/Disagree/Strongly Disagree	Agree
16. During the past month, I drank alcohol, used marijuana, cocaine, or street drugs.	Strongly Agree/Agree/Don't Know/Disagree/Strongly Disagree	Disagree
17. I feel able to say "no" to a friend who wants me to do something that I shouldn't do.	Strongly Agree/Agree/Don't Know/Disagree/Strongly Disagree	Agree
18. Not having sex is the best way to avoid getting HIV or STI.	Strongly Agree/Agree/Don't Know/Disagree/Strongly Disagree	Agree
19. I know who/where to go to in my community when I need help or need to talk.	Strongly Agree/Agree/Don't Know/Disagree/Strongly Disagree	Agree
20. I plan to wait until I am older to have sex.	Strongly Agree/Agree/Don't Know/Disagree/Strongly Disagree	Agree

The pre-test and post-test also collected information on participant race/ethnicity and age. The pre-test and post-test are administered during the first and last sessions of the series, respectively. The facilitator is instructed to take about 10 minutes to administer the test. Each student receives a copy of the test, and the facilitator reads each question aloud, with students following along. A copy of the test and instructions for facilitator administration can be found in Appendices 2 and 3. Beginning with the Fall 2013 series at John Harris HS, an employee of FHCCP was in attendance at the first and last sessions to proctor the tests.

Section 3.4: Analysis

3.4.1: Demographic Data

All races and ethnicities listed by students in their test responses were retained for analysis. Race data were dichotomized for two variables. For one variable, all students who listed "Black/African-American," whether alone or together with another race/ethnicity, were compared to those who did not. In the second variable, the same was done for students who listed "Hispanic/Latino."

Age data was dichotomized in three variables: one compared those under the age of 15 with those above, and the other two did the same with those under 16 and those above, and those under 17 and those above.

3.4.2: Intervention Exposure

Intervention exposure was measured by characterizing sessions as "successful" or "unsuccessful." A "successful" session was defined by the following characteristics:

- The student was in attendance
- "All" or "some" of the students participated
- "All" or "some" of the students were interested
- At least 3 students were in attendance

- At least 60% of activities were conducted either as described or with changes

Both the number of sessions attended (whether successful or not) and the number of successful sessions attended were dichotomized for analysis into “7 or fewer” sessions and “8 or more sessions.”

In analysis, it was necessary to divide up the sessions beyond the seven described by the curriculum. Some sites, most notably John Harris, split one session over two or even three dates. This meant that attendance data was gathered for each date; many students were therefore present at one part of “Session 4” but not the other(s). To reflect this issue, each session was divided into parts, showing the largest number of parts into which it was divided by any site. For example, three of the John Harris classes conducted Session 4 over 3 separate dates. In the data, Sessions 4a, 4b, and 4c were shown, with different attendance data for each. To allow comparison with sites that did not split up sessions in this way, data for that particular session was copied across all three spaces. For example, the New Beginnings classes did not split up any sessions. Therefore the data for the New Beginnings Session 4 was copied across these three spaces. Using this method, there were a total of 14 partial sessions.

Table 3: Division of Sessions for Analysis

Original session	Sessions for analysis
Session 1	1a
	1b
Session 2	2a
	2b
Session 3	3a
	3b
	3c
Session 4	4a
	4b
	4c
Session 5	5
Session 6	6
Session 7	7a
	7b

3.4.3: Pre-Test and Post-Test Responses

Several different analysis strategies were used. First, in analysis of post-test scores, response to Likert Scale items were dichotomized as "correct" or "incorrect"; if the desired response to the item was "strongly agree," both "strongly agree" and "agree" were coded as "correct" while "strongly disagree" and "disagree" were coded as "incorrect." Responses of "don't know" were coded as "incorrect" as well. Responses to Yes/No items were coded similarly; if the desired answer was "yes," responses of "yes" were coded as correct and responses of "no" or "don't know" were coded as "incorrect."

Reponses were also coded numerically, Likert Scale items being coded from 1 to 5. If the desired response was "strongly agree", "strongly agree" was coded as 5, "agree" was coded as 4, "don't know" was coded as 3, "disagree" was coded as 2, and "strongly disagree" was coded as 1. For Yes/No items, if "yes" was the desired response, "yes" was coded as 5, "don't know" as 3, and "no" as 1. This allowed the determination of means across groups of participants for individual items, groups of items, and for all items together.

Questions were analyzed using these two methods on an individual basis. Questions were also grouped thematically in 9 different composite items. These groupings were based on topic area: STIs, pregnancy, drugs, and condoms; and on competency area: knowledge, attitudes, self-efficacy, behaviors, and intentions. The details of these composite items are shown in the table below.

Table 4: Composite Items for Analysis

Composite Item	Questions Included	Points Possible
STIs	2, 3, 4, 7, 15	20
Pregnancy	1, 9	8
Drugs	6, 11	8
Condoms	7, 12, 14	12
Knowledge	1, 2, 3, 4, 6, 7, 15	28
Attitudes	5, 9, 10, 11, 13, 18	24
Self-efficacy	12, 17, 19	12
Behaviors	8, 16	8
Intentions	14, 20	8

Change was pre-test to post-test was also analyzed in a number of ways. Using the numeric coding described above, the pre-test response was subtracted from the post-test response; higher numbers indicate a more desirable change, while negative numbers would indicate undesirable change. The change from pre-test to post-test was also dichotomized, with the result of the subtraction described above coded as “positive change” if above 0, and as “no positive change” if equal to or less than 0.

These approaches allowed both determination of means and comparison using chi-square analysis, and determination of measures of association between exposure and outcomes using dichotomized data.

Section 3.5: Ethics

3.5.1: Consent

Passive consent procedures were used to obtain parental consent. An opt-out letter was sent home with students prior to the first session describing the O2T program and giving parents to opportunity to deny

permission for their child to take part. If the opt-out form was not received by the deadline, parental consent was assumed. A copy of the parental opt-out letter and form can be found in Appendix 4.

Participants were informed that their participation in the program and their completion of the pre-test and post-test were voluntary and that they could withdraw or decline to participate at any time. A few students availed themselves of this option, but in general, students who parents did not decline permission agreed to participate themselves. In addition, in all cases, the O2T program was provided within the context of a specific high-school class or summer program in which students had voluntarily elected to enroll while knowing that O2T content would be involved.

3.5.2: Confidentiality

Participants were assigned identification numbers and these were used to track test responses. Facilitators recorded data including names, dates of birth, and session attendance, which also contained identification numbers, in separate files from those containing pre-test and post-test responses. In no files were identifying details directly linked with test responses. Facilitators promptly destroyed paper records such as copies of tests after entering data into electronic files. All files were kept on a secure password-protected server. Good data-protection and confidentiality procedures were practiced by program staff, including avoiding creating hard copies of files or data, using password protection on personal and work computers, and never leaving windows containing identifying information open and unattended in the office.

Chapter 4: Results

Section 4.1: Characteristics of Participants

The data being considered consists of 213 students participating in 11 different series. All but one of the series was conducted in the Harrisburg metropolitan area by the Cumberland-Perry Tapestry of Health, a direct service site of the FHCCP. The series are divided into three cohorts. The first cohort, called "New Beginnings", consists of six series. New Beginnings is a summer youth enrichment day program that contracted with TOH to host these series at three different locations – John Harris, Susquehanna Township, and Steelton. The first series were hosted at these sites in Summer 2012, between mid-June and mid-July ("New Beginnings John Harris, Susq. Twp, and Steelton 2012"). Three series were hosted at the same sites the following summer, between late June and late July ("New Beginnings John Harris, Susq. Twp, and Steelton 2013").

Another four series were hosted by John Harris High School during the regular school year, as part of several different parenting and life skills classes. The first of these classes was conducted in the 2013 Spring Semester, between late January and mid-April ("John Harris Spring 2013"). There were 31 participants recorded in this class. The other three classes were conducted during the 2013 Fall Semester, between early October and early December ("John Harris Fall 2013 A, B, and C").

The one remaining series was conducted by another partner; the Families United Network, located in Williamsport, PA. This series was conducted during Summer 2013 between early July and mid-August at Ashler Manor, a residential program for adolescent females referred by Children and Youth Services or Juvenile Probation ("Families United").

The table below shows the distribution of participants across the eleven series and three cohorts.

Table 5: O2T participants by site and class

Site	Class	No. students	Provider
John Harris HS	Spring 2013	31	CP TOH
John Harris HS	Fall 2013 A	6	CP TOH
John Harris HS	Fall 2013 B	32	CP TOH
John Harris HS	Fall 2013 C	32	CP TOH
Total John Harris Cohort		101	
New Beginnings	John Harris Summer 2012	10	CP TOH
New Beginnings	Susq. Twp Summer 2012	19	CP TOH
New Beginnings	Steelton Summer 2012	5	CP TOH
New Beginnings	John Harris Summer 2013	11	CP TOH
New Beginnings	Susq. Twp Summer 2013	19	CP TOH
New Beginnings	Steelton Summer 2013	10	CP TOH
Total New Beginnings Cohort		74	
Ashler Manor	Summer 2013	38	Families United

4.1.1: Age

Of the 213 participants recorded, 178 had information on age. Of these, 159 had full birthdates, 4 had only birth years, and 15 had only ages as of the first session. For those with only a birth year, a birthdate of January 1 was assumed. The most exact age possible was used. The average age of participants was 16.5, ranging from 12.0 to 20.9. On average, the New Beginnings and Families United cohorts were younger than the John Harris cohort, with mean ages of 15.8 and 15.7 versus 17.1, respectively.

The youngest student at John Harris was just over the age of 15. The New Beginnings cohort had only student just over age 19; other than this student, no participants were 18 or older. This cohort also had the highest number of young students; 14 participants were under the age of 15.

Overall, 77% of students were aged 15- 17; with 8% aged under 15 and 14% aged 18 or over.

Figure 7: Age Distribution of O2T Participants

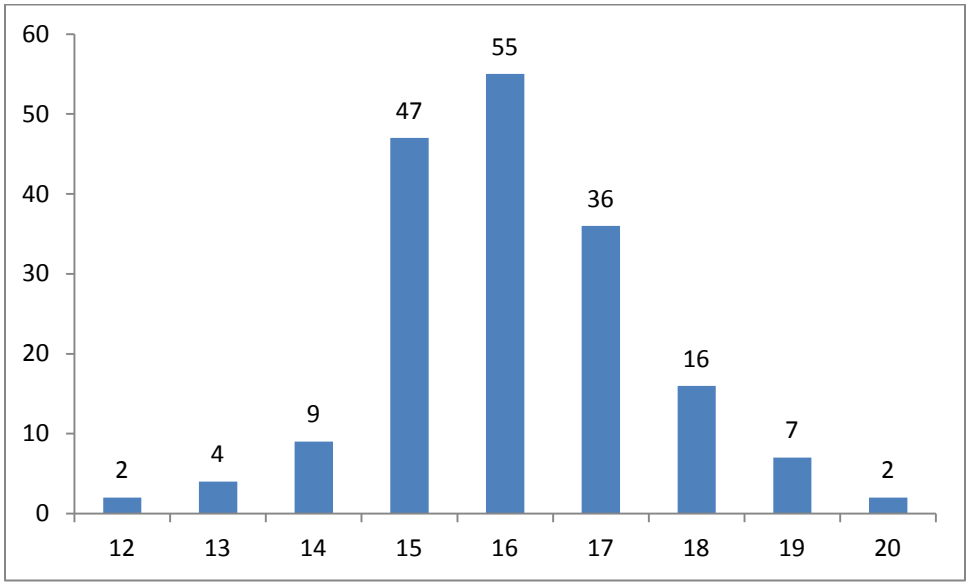


Table 6: Age of O2T Participants by Site

	Number of students	Minimum	Median	Mean	Maximum
Families United	15	12	16	15.7	18
JH	96	15.1	17.0	17.1	20.9
NB	67	12.4	15.9	15.9	19.1

4.1.2: Race/Ethnicity

Less than half of participants had race information recorded; 102 out of 213. The largest group of students with missing data on race and ethnicity occurred in the Summer 2012 New Beginnings cohort. All 34 of these students lacked data on race. Students in the other cohorts lacked data on race if they were absent at the first session of the series.

Only 6 of these students considered themselves White alone; 50 considered themselves Black alone, 20 considered themselves Hispanic alone, 8 considered themselves Asian or Pacific Islander alone, and 3 considered themselves Native American alone.

A total of 29 students, or 28.4%, claimed Hispanic ethnicity either alone or in combination, while 55, or 56.8%, considered themselves Black either alone or in combination.

In the Summer 2013 cohort, the majority of students – 27 out of 40, or 67.5% – considered themselves Black either alone or in combination with another race or ethnicity. Nine students considered themselves Hispanic either alone or in combination, with five indicating other races.

The John Harris cohort is also largely Black and Hispanic, with 22 out of 47 students (46.8%) considering themselves Black and 16 considering themselves Hispanic (34%). This cohort had the only significant population of students who considered themselves Asian – 6 out of 47, or 12.8%. An additional 2 students considered themselves White. Only 15 of the 38 students within the Families United cohort had data on race. Four students – or 26.7%, the most of any cohort – considered themselves White, with five considering themselves Black and six Hispanic.

Figure 8: Race and Ethnicity of O2T Participants

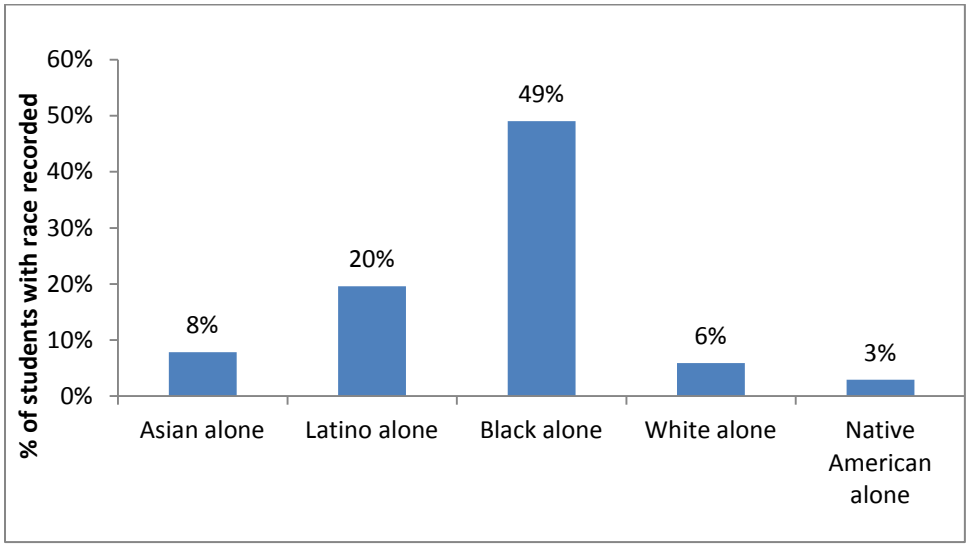


Table 7: Race and Ethnicity of O2T Participants

Race/Ethnicity	Number of students	% of those with race recorded
Asian alone	8	7.84%
Asian alone or in combination	9	8.82%
Latino alone	20	19.61%
Latino alone or in combination	29	28.43%
Black alone	50	49.02%
Black alone or in combination	55	53.92%
White alone	6	5.88%
White alone or in combination	15	14.71%
Native American alone	3	2.94%
Native American alone or in combination	6	5.88%
Total students with race recorded	**102**	**100%**

4.1.3: Sex

All but one of the 213 participants had gender information recorded; 58% of these were female and 42% were male. The gender distribution varied by site, with all of the 38 participants in the Families United cohort being female, compared to 48.5% of participants in the John Harris cohort and exactly 50% of the New Beginnings cohort.

Section 4.2: Intervention Exposure

4.2.1: Fidelity Monitoring

Each series consists of 37 activities across 7 sessions. Eleven full series were conducted, giving a total of 407 activities performed. Of these, 190, or 46.7%, were performed as described, 125 or 30.7% were performed with changes, and 85 or 20.9% were not performed at all.

The activities most often omitted include Activity 2.4 (omitted 10 of 11 times) and activities 6.3 and 6.4 (each omitted 8 of 11 times). The activities changed most frequently include Activity 1.5 (changed 10 of 11 times) and activity 1.4 (changed 9 of 11 times). The activities conducted with the most fidelity include activities 1.1 and 1.3, each conducted as described every time, and activities 2.3 and 5.3, each conducted as described 10 of 11 times.

One activity – 6.5 – was never conducted as described. It was omitted 7 times and changed 4 times. Activities 1.5, 3.1, and 5.1 were each only conducted as described once. Activities 1.4, 3.2, 6.3, 6.4, and 7.4 were all conducted as described only twice.

Table 8: Activity Implementation Fidelity by Site

Site	JH	JH	JH	JH	JH	Families United	NB	NB	NB	NB	NB	NB	NB
Class	Fall 2013A	Fall 2013B	Fall 2013C	Spring 2013	Avg	Ashler Manor	JH 2012	Susq 2012	Steel 2012	JH 2013	Steel 2013	Susq. 2013	Avg
As described	24	22	24	15	21.25	22	20	17	9	14	10	13	13.83
With changes	8	8	7	14	9.25	14	11	13	11	10	14	15	12.33
Not conducted	5	6	6	8	6.25	1	4	5	15	13	13	9	9.83

Table 9: Implementation Fidelity of All Activities

Activity	Conducted as described	Conducted with changes	Not conducted
Act. 1.1	11	0	0
Act. 1.2	6	5	0
Act. 1.3	11	0	0
Act. 1.4	2	9	0
Act. 1.5	1	10	0
Act. 1.6	8	3	0
Act. 2.1	8	1	2
Act. 2.2	5	6	0
Act. 2.3	10	0	0
Act. 2.4	1	0	10
Act. 2.5	4	5	2
Act. 2.6	7	4	0
Act 3.1	1	5	5
Act. 3.2	2	6	3
Act. 3.3	5	6	0
Act. 3.4	6	5	0
Act 4.1	7	4	0
Act. 4.2	6	5	0
Act. 4.3	3	6	2
Act. 4.4	4	3	4
Act. 4.5	4	6	1
Act. 5.1	1	4	6
Act. 5.2	3	1	7
Act. 5.3	10	0	1
Act. 5.4	8	3	0
Act. 5.5	6	2	3
Act. 6.1	8	3	0
Act. 6.2	9	1	1
Act. 6.3	2	1	8
Act. 6.4	2	1	8
Act. 6.5	0	4	7
Act. 6.6	8	1	2
Act. 7.1	7	1	3
Act. 7.2	5	1	2
Act. 7.3	3	7	1
Act. 7.4	2	5	4
Act. 7.5	4	1	3
Total	**190**	**125**	**85**
Percent	**46.7%**	**30.7%**	**20.9%**

4.2.2: Attendance and Participation

On average, each student attended 7.6 sessions out of the 14 possible. Twelve of the 213 students recorded, or 5.6%, never attended a single session, while 25, or 11.7% attended all 14. The average number of attendees at any particular session was 11.2. The average number of attendees was fairly constant across sessions, ranging from 10.1 at Session 7b to 12.6 at Session 1a. However, the average number of attendees varied significantly across sites. The Summer 2012 New Beginnings Steelton series had an average of only 2.4 attendees each session. However, because the enrollment in the class was only 5 students, the average percentage of enrolled students present was 47.1%, which is higher than the average at Ashler Manor (45.1%) and not far below that of other classes. The average proportion of enrolled students present at each session was 57.8%. This was lowest at Ashler Manor, as mentioned, and highest at the Summer 2012 New Beginnings John Harris site, with an average of 7.4, or 73.6% of enrolled students present at each session.

Figure 9: Number of sessions attended as percent of all students

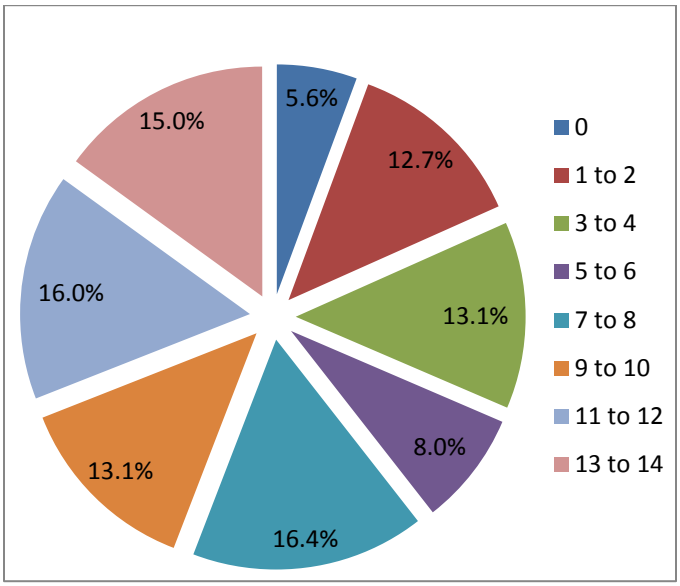


Table 10: Number of Sessions Attended, Frequency and as Percent of all Students

Number sessions attended	Number students	Percent
0	12	5.6%
1	4	1.9%
2	23	10.8%
3	6	2.8%
4	22	10.3%
5	10	4.7%
6	7	3.3%
7	9	4.2%
8	26	12.2%
9	14	6.6%
10	14	6.6%
11	16	7.5%
12	18	8.5%
13	7	3.3%
14	25	11.7%

Table 11: Number of students in attendance at each session by site and class

Site	Class	Total enrolled	Sess. 1a	Sess. 1b	Sess. 2a	Sess. 2b	Sess. 3a	Sess. 3b	Sess. 3c	Sess. 4a	Sess. 4b	Sess. 4c	Sess. 5	Sess. 6	Sess. 7a	Sess. 7b	Average
John Harris HS	Spring 2013	31	22	22	21	20	9	9	9	14	14	14	14	14	21	16	**15.6**
John Harris HS	Fall 2013 1	6	4	3	2	4	3	1	3	2	2	3	4	4	4	4	**3.1**
John Harris HS	Fall 2013 2	32	24	21	16	15	20	18	21	21	19	20	21	22	22	19	**19.9**
John Harris HS	Fall 2013 3	32	22	21	21	15	19	19	20	24	24	20	19	22	22	19	**20.5**
New Beginnings	John Harris Summer 2012	10	7	7	10	10	10	10	10	7	7	7	4	2	6	6	**7.4**
New Beginnings	Susq. Twp Summer 2012	19	13	13	12	12	12	12	12	11	11	11	11	12	13	13	**12.0**
New Beginnings	Steelton Summer 2012	5	3	3	4	4	4	4	4	1	1	1	1	1	1	1	**2.4**
New Beginnings	John Harris Summer 2013	11	7	7	7	7	7	7	7	7	7	7	6	6	6	6	**6.7**
New Beginnings	Susq. Twp Summer 2013	19	12	12	12	12	12	12	12	12	12	12	12	12	12	12	**12.0**
New Beginnings	Steelton Summer 2013	10	9	9	9	9	7	7	7	7	7	7	5	5	0	0	**6.3**
Ashler Manor	Summer 2013	38	16	16	18	18	19	19	18	18	18	18	16	16	15	15	**17.1**
Average		**19.4**	**12.6**	**12.2**	**12.0**	**11.5**	**11.1**	**10.7**	**11.2**	**11.3**	**11.1**	**10.9**	**10.3**	**10.5**	**11.1**	**10.1**	**11.2**

4.2.3: Successful Sessions

A "successful" session was defined by the following characteristics:

- The student was in attendance

- “All” or “some” of the students participated
- “All” or “some” of the students were interested
- At least 3 students were in attendance
- At least 60% of activities were conducted either as described or with changes

The average number of successful sessions attended was 6.4. Fourteen, or 6.6% of recorded students, did not attend a single successful session. Twelve of these attended no sessions at all (see above), while the other two attended only unsuccessful sessions. Only six, or 2.8% of students, attended 14 successful sessions out of a possible 14.

Table 12: Number of successful sessions attended, frequency and as percent of all students

Number successful sessions attended	Number students	Percent
0	14	6.6%
1	11	5.2%
2	37	17.4%
3	13	6.1%
4	11	5.2%
5	10	4.7%
6	5	2.3%
7	11	5.2%
8	22	10.3%
9	20	9.4%
10	19	8.9%
11	9	4.2%
12	12	5.6%
13	13	6.1%
14	6	2.8%

Figure 10: Number of successful sessions attended as percent of all students

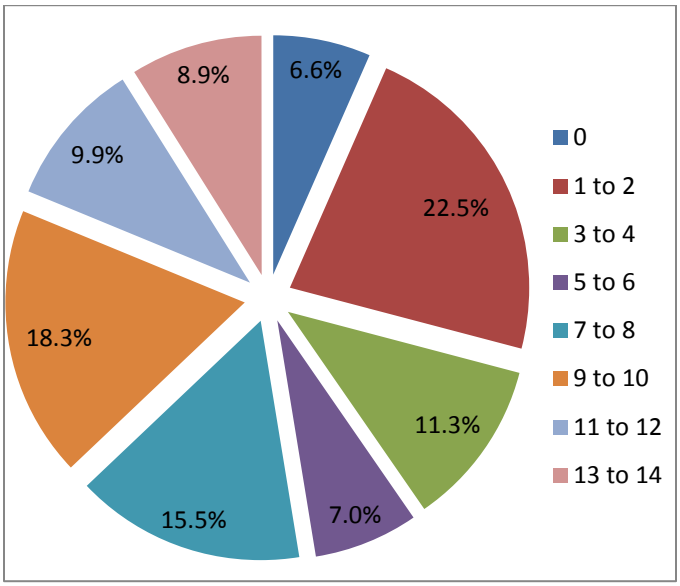


Table 13: Successful and unsuccessful sessions by site and class, number & percent of all sessions

Site	Class	Successful		Unsuccessful	
John Harris	Spring 2013	12	86%	2	14%
John Harris	Fall 2013 A	10	71%	4	29%
John Harris	Fall 2013 B	13	93%	1	7%
John Harris	Fall 2013 C	13	93%	1	7%
New Beginnings	JH 2012	8	57%	6	43%
New Beginnings	Susq. 2012	11	79%	3	21%
New Beginnings	Steel. 2012	4	29%	10	71%
New Beginnings	JH 2013	10	71%	4	29%
New Beginnings	Susq. 2013	11	79%	3	21%
New Beginnings	Steel. 2013	9	64%	5	36%
Families United	Ashler Manor	14	100%	0	0%

Section 4.3: Pre-Test Scores

4.3.1: Scores and Percent Correct

Among the 213 recorded participants, 133 took the pre-test. On average, the total score on the pre-test was 60.5 points out of a possible 80. Three participants, or 2.3%, scored less than 40/80. Fourteen (10.5%) scored between 40 and 50; 38 (28.6%) scored between 50 and 60; 52 (39.1%) scored between

60 and 70; and 26 (19.5%) scored over 70. There were no perfect scores – the closest was 79/80. The lowest score was 34/80.

On average, participants answered 14.7 questions correctly out of 20, or 73.4%. The lowest number of correct answers was 2, and the highest was 20. Only two participants answered all 20 questions correctly; however another ten answered 19 correctly and 15 answered 18 correctly. Overall, 79 participants, or 59.4% of those who took the pre-test, answered 15 or more questions correctly; 119, or 89.4%, answered 10 or more questions correctly.

Scores on the pre-test varied between sites. The Families United site in Williamsport had the highest scores, with an average of 16.1 questions correct out of 20 and a score of 63.4 points out of 80. New Beginnings had the lowest scores, with 13.7 questions correct and a score of 58.6 out of 80. John Harris fell in the middle, with 15.1 questions correct and an average score of 61.3 out of 80.

Table 14: Pre-Test, percent answering question correctly and average score on question

Question	Pre-Test % Correct	Score
1. Can a girl get pregnant the first time she has sex?	85.8%	3.5373
2. Can you get STI's from having oral sex?	73.7%	3.3985
3. Can you get HIV from kissing someone?	54.1%	2.5414
4. Can a person have an STI and not know it?	79.7%	3.5639
5. Can you say "no" to sex if you love that person?	82.0%	3.4436
6. Can some drugs kill you, even if you use them one time?	85.0%	3.6241
7. Can wearing a condom keep you from getting an STI or HIV?	47.7%	2.1667
8. In the past month I have had sex (vaginal, anal, or oral.)	69.5%	2.771
9. Getting pregnant or getting someone pregnant at my age would make it more difficult to achieve my goals.	78.9%	3.1504
10. Having sex would be a good way to keep a boyfriend or girlfriend.	68.4%	2.7519
11. I am too young to do drugs or drink alcohol.	66.4%	2.6641
12. I know how to use a condom.	78.2%	3.1805
13. I feel that I have a number of good qualities.	85.6%	3.2955
14. I would insist that my partner or I use a condom every time we had sex.	75.0%	3.0682
15. If a person has unprotected sex one time, s/he can get an STI or HIV.	82.6%	3.3106
16. During the past month, I drank alcohol, used marijuana, cocaine, or street drugs.	68.9%	2.8788
17. I feel able to say "no" to a friend who wants me to do something that I shouldn't do.	85.6%	3.3106
18. Not having sex is the best way to avoid getting HIV or STI.	74.8%	3.0763
19. I know who/where to go to in my community when I need help or need to talk.	78.8%	3.053
20. I plan to wait until I am older to have sex.	32.6%	1.7045
Total	**73.4%**	**60.5**

The answers with the highest scores included question 6 (3.62), question 4 (3.56), and question 1 (3.54). The answers with the lowest scores included question 20 (1.71), question 7 (2.17), and question 3 (2.54).

The answers with the highest percent correct included question 1 (85.8%), question 17 (85.6%), and question 13 (85.6%). The answers with the lowest percent correct included question 20 (32.6%), question 7 (47.7%), and question 3 (54.1%).

4.3.2: Distribution of Response Values

Questions differed in how responses were distributed over the spectrum of correct and incorrect answers. Several questions saw a high percentage of students responding “don’t know.” The highest among these were questions 2 (22.6%), 3 and 4 (18.8% each), and 15 (14.4%). The questions with the lowest percentage of students unsure of the correct answer were questions 17 (4.5%), 11 (4.6%), and 1 (5.2%). For questions 8 through 20, which used a Likert-type scale for responses, questions also differed by the percentage of students choosing one of the “strong” response values – either “strongly agree” or “strongly disagree”. These ranged from 40.1% (question 20) to 65.9% (question 17).

Table 15: Distribution of Responses to Pre-test questions

Question	Yes		Don't Know		No
1	85.80%		5.2%		9.0%
2	73.70%		22.6%		3.8%
3	27.10%		18.8%		54.1%
4	79.70%		18.8%		1.5%
5	82.00%		8.3%		9.8%
6	85.00%		11.3%		3.8%
7	47.70%		12.9%		39.4%
Question	**Strongly Agree**	**Agree**	**Don't Know**	**Disagree**	**Strongly Disagree**
8	12.2%	12.2%	6.1%	25.2%	44.3%
9	55.6%	23.3%	8.3%	6.0%	6.8%
10	7.5%	14.3%	9.8%	32.3%	36.1%
11	38.2%	28.2%	4.6%	19.8%	9.2%
12	52.6%	25.6%	14.3%	2.3%	5.3%
13	51.5%	34.1%	9.1%	3.0%	2.3%
14	46.2%	28.8%	11.4%	12.9%	0.8%
15	53.8%	28.8%	14.4%	0.8%	2.3%
16	9.8%	13.6%	7.6%	16.7%	52.3%
17	60.6%	25.0%	4.5%	4.5%	5.3%
18	54.2%	20.6%	9.2%	10.7%	5.3%
19	43.9%	34.8%	8.3%	8.3%	4.5%
20	15.9%	16.7%	13.6%	29.5%	24.2%

4.3.3: Composite Items

The percentage of questions answered correctly within each composite item varied from 53.8% to 83%. The composite item with the lowest score (53.8%) was the intentions item; its low score largely due to the frequency of students responding “incorrectly” to question 20 (indicating that they did not plan to

wait to have sex). The highest percent of questions correct was found for the pregnancy composite item, reflecting the high percentage of students responding correctly to questions 1 and 9.

Table 16: Pre-test composite items: percent of questions answered correctly

Composite item	% questions correct
STIs	67.73%
Pregnancy	82.95%
Drugs	75.76%
Condoms	67.17%
Attitudes	76.14%
Behavior	68.94%
Intentions	53.79%
Knowledge	72.94%
Self-efficacy	81.06%

4.3.4: Demographic Variation

Some differences were seen in pre-test scores according to demographic variables. Black students were 2.5 times more likely than students of other races and ethnicities to respond correctly to question 2, which asked "can you get STIs from having oral sex?" (p=0.049). Black students were also 3.1 times more likely to report using drugs or alcohol within the last month (p=-0.024). Latino students were 12.6 times more likely to answer question 1 ("Can a girl get pregnant the first time she has sex?") correctly than students of other races and ethnicities (p=0.008).

Female students were 3.2 times more likely than male students to respond correctly to question 10, which asked students to agree or disagree with the statement "Having sex would be a good way to keep a girlfriend or boyfriend" (p=0.002).

Students under the age of 15 were less likely to answer questions 1, 2, 12, 13, 14, and 15 correctly. They were also less likely to achieve a score of more than 75% on the knowledge, STIs, pregnancy, and self-efficacy composite items. However, they were more likely to indicate that they were planning on waiting to have sex (question 20).

Table 17: Relationship between age of under 15 years and scores on pre-test items

Question/Index item	Odds Ratio	95% Confidence Interval	p-value
1	0.0306	(0.0057-0.1643)	<0.0001
2	0.1543	(0.0363-0.6557)	0.0108
12	0.115	(0.0268-0.4943)	0.0037
13	0.1752	(0.0423-0.7261)	0.0253
14	0.2383	(0.0599-0.9480)	0.0443
15	0.1385	(0.0339-0.5653)	0.0086
20	8.3611	(1.6564-42.2049)	0.0058
Knowledge	0.121	(0.0147-0.9968)	0.0218
STIs	0.1026	(0.0124-0.8457)	0.0121
Pregnancy	0.0932	(0.0184-0.4730)	0.0021
Self-efficacy	0.1612	(0.0321-0.8098)	0.0178

A similar pattern was seen for students under 16 years old. Like those under 15, these students were less likely to answer questions 1, 2, 12, 13, and 15 correctly, however the effect of age on the percentage of students answering correctly was less strong. The relationship between younger age and answering question 14 incorrectly was no longer significant when divided at this point, but the relationship between younger age and answering question 5 incorrectly became significant. Students under the age of 16 were also more likely to indicate waiting to have sex (question 20), but this association was not as strong as for those under age 15. The same was true for composite items dealing with pregnancy and self-efficacy; the relationships between younger age and percent of questions in each composite item was still significant in this case, but was less strong. The relationships between younger age and lower percent of questions correct on the knowledge and STI composite items was no longer significant in this case.

Table 18: Relationship between age under 16 years and scores on pre-test items

Question/Index item	Odds Ratio	95% Confidence Interval	p-value
1	0.1724	(0.0602-0.4934)	0.0007
2	0.4084	(0.1839-0.9070)	0.0228
5	0.4211	(0.1711-1.0361)	0.0491
12	0.2508	(0.1062-0.5924)	0.0014
13	0.3038	(0.1121-0.8235)	0.0171
15	0.3097	(0.1229-0.7801)	0.0114
20	3.1429	(1.4576-6.7767)	0.0029
Pregnancy	0.4741	(0.2160-1.0404)	0.0484
Self-efficacy	0.4333	(0.2061-0.9111)	0.0212

For students under the age of 17, the relationship between younger age and the percentage of students answering questions 12 and 13 incorrectly persisted, as did the percentage of students indicating that they planned to wait to have sex (question 20). Interestingly, while all three of these relationships were less strong than when applied to students under the age of 15, they were stronger than when applied to students under the age of 16. Furthermore, the relationship between age and scores on the pregnancy and self-efficacy composite items was no longer significant, but the relationships between age and scores on the knowledge and STI composite items *was* significant, though it was not among students under 16. The association for each of these items was less strong than it was for those under 15.

Table 19: Relationship between age under 17 years and scores on pre-test items

Question/ Composite item	Odds Ratio	95% Confidence Interval	p-value
12	0.2233	(0.0724-0.6882)	0.0038
13	0.1752	(0.0386-0.7954)	0.0092
20	3.325	(1.3843-7.9866)	0.0042
STIs	0.4312	(0.2053-0.9058)	0.0193

Knowledge	0.45	(0.2162-0.9365)	0.0242

Section 4.4: Post-Test Scores

4.4.1: Scores and Percent Correct

Among the 213 recorded participants, 93 took the post-test. On average, the total score on the post-test was 62.8 points out of a possible 80. Two participants, or 2.2%, scored less than 40/80. Four (4.3%) scored between 40 and 50; 20 (21.5%) scored between 50 and 60; 48 (51.6%) scored between 60 and 70; and 19 (20.4%) scored over 70. There were no perfect scores – the closest was 77/80. The lowest score was 38/80.

On average, participants answered 15.6 questions correctly out of 20, or 77.8%. The lowest number of correct answers was 2, and the highest was 20. Only one participant answered all 20 questions correctly; however another11 answered 19 correctly and 9 answered 18 correctly. Overall, 34 participants, or 36.6% of those who took the post-test, answered 15 or more questions correctly; 88, or 94.6%, answered 10 or more questions correctly.

Scores on the pre-test varied between sites. The Families United site in Williamsport had the highest scores, with an average of 16.3 questions correct out of 20 and a score of 64.6 points out of 80. New Beginnings had the lowest scores, with 14.8 questions correct and a score of 62.1 out of 80. John Harris fell in the middle, with 15.8 questions correct and an average score of 62.8 out of 80.

Table 20: Post-Test, percent answering question correctly and average score on question

Question	% Correct	Score
1. Can a girl get pregnant the first time she has sex?	97.8%	3.9355
2. Can you get STI's from having oral sex?	83.0%	3.5745
3. Can you get HIV from kissing someone?	58.1%	2.5806
4. Can a person have an STI and not know it?	81.7%	3.5699
5. Can you say "no" to sex if you love that person?	88.2%	3.6129

6. Can some drugs kill you, even if you use them one time?	88.0%	3.6957
7. Can wearing a condom keep you from getting an STI or HIV?	48.4%	2.1538
8. In the past month I have had sex (vaginal, anal, or oral.)	76.3%	2.9785
9. Getting pregnant or getting someone pregnant at my age would make it more difficult to achieve my goals.	80.6%	3.1935
10. Having sex would be a good way to keep a boyfriend or girlfriend.	69.9%	2.7527
11. I am too young to do drugs or drink alcohol.	72.0%	2.8495
12. I know how to use a condom.	78.3%	3.1522
13. I feel that I have a number of good qualities.	90.3%	3.3763
14. I would insist that my partner or I use a condom every time we had sex.	83.5%	3.3297
15. If a person has unprotected sex one time, s/he can get an STI or HIV.	90.1%	3.3516
16. During the past month, I drank alcohol, used marijuana, cocaine, or street drugs.	78.0%	3.1758
17. I feel able to say "no" to a friend who wants me to do something that I shouldn't do.	87.0%	3.3152
18. Not having sex is the best way to avoid getting HIV or STI.	81.5%	3.2174
19. I know who/where to go to in my community when I need help or need to talk.	85.9%	3.0978
20. I plan to wait until I am older to have sex.	38.0%	1.9239
Total	**77.8%**	**62.8**

4.3.2: Distribution of Response Values

Questions differed in how responses were distributed over the spectrum of correct and incorrect answers. Several questions saw a high percentage of students responding “don’t know.” The highest among these were questions 4 (15.1%), 3 (12.9%), and 2 (12.8%). The questions with the lowest percentage of students unsure of the correct answer were questions 1 (1.1%), 16 and 17 (both 2.2%), and 11 (3.2%). For questions 8 through 20, which used a Likert-type scale for responses, questions also differed by the percentage of students choosing one of the “strong” response values – either “strongly agree” or “strongly disagree”. These ranged from 39.1% (question 20) to 72.5% (question 16).

Table 21: Distribution of Responses to Post-test Questions

Question	Yes		Don't Know		No
1	97.8%		1.1%		1.1%
2	83.0%		12.8%		4.3%
3	29.0%		12.9%		58.1%
4	81.7%		15.1%		3.2%
5	88.2%		4.3%		7.5%
6	88.0%		8.7%		3.3%
7	48.4%		11.0%		40.7%
Question	**Strongly Agree**	**Agree**	**Don't Know**	**Disagree**	**Strongly Disagree**
8	11.8%	8.6%	3.2%	22.6%	53.8%
9	60.2%	20.4%	3.2%	10.8%	5.4%
10	35.5%	34.4%	6.5%	17.2%	6.5%
11	44.1%	28.0%	3.2%	18.3%	6.5%
12	51.1%	27.2%	12.0%	5.4%	4.3%
13	52.7%	37.6%	5.4%	3.2%	1.1%
14	54.9%	28.6%	12.1%	3.3%	1.1%
15	53.8%	36.3%	4.4%	2.2%	3.3%
16	6.6%	13.2%	2.2%	12.1%	65.9%
17	60.9%	26.1%	3.3%	3.3%	6.5%
18	59.8%	21.7%	5.4%	6.5%	6.5%
19	42.4%	43.5%	2.2%	5.4%	6.5%
20	21.7%	16.3%	12.0%	32.6%	17.4%

4.3.3: Composite Items

The percentage of questions answered correctly within each composite item varied from 53.8% to 83%. The composite item with the lowest score (53.8%) was the intentions item; its low score largely due to the frequency of students responding “incorrectly” to question 20 (indicating that they did not plan to

wait to have sex). The highest percent of questions correct was found for the pregnancy composite item, reflecting the high percentage of students responding correctly to questions 1 and 9.

Table 22: Post-test composite items; percent of questions answered correctly

Composite item	% questions correct
STI	71.61%
pregnancy	89.25%
drugs	79.57%
condoms	68.82%
attitudes	80.29%
behavior	76.34%
intentions	59.68%
knowledge	77.57%
self-efficacy	82.80%

Section 4.5: Change from Pre- and Post-Tests

4.5.1: Missing Tests

Only 74 participants had both a pre-test and a post-test recorded. Fifty-nine students took the pre-test and not the post-test, and 19 took the post-test but not the pre-test. The missing pre-tests were due entirely to students either being absent at the first session or not enrolling in the class until after the first session. Most of the missing post-tests were due to absences or withdrawal from the class before the end of the series, but others were due to additional factors. The last session of the Summer 2013 series of the Steelton New Beginnings program was cancelled without warning and not made up, meaning that all ten students enrolled in this series were not exposed to the content of the final session and did not complete post-tests. The post-test data from the Summer 2013 John Harris New Beginnings series is missing, and either the six participants present that day never completed their post-tests, or the data was never recorded. Sixty-one students took neither test; five of these are accounted for by the

problems with the 2013 New Beginnings post-tests, but the remainder were absent at both the first and last sessions. Twelve students are recorded as never attending a single session.

Missing Pre-test

While most of those students without a pre-test had no race/ethnicity information recorded, 11 of the 80 students without a pre-test did have this information. According to this very limited dataset, Black students were more likely to be missing a pre-test. Among those with race information recorded, 18.5% of Black students missed the pre-test compared to 2.4% of students of other races and ethnicities ($p=0.014$). There were no differences in mean age or in gender between those with a pre-test and those without.

Those without a pre-test attended significantly fewer sessions than those with a pre-test; 8.8 for those with a pre-test compared to 5.6 for those without ($p<0.0001$). However, when excluding those who also missed the post-test, there was no significant difference in the number of session attended between those that missed the pre-test but took the post-test and those who took both tests ($p=0.14$).

There were no significant differences in the percent of questions correct on the post-test between those who missed the pre-test and those who didn't ($p=0.089$). The mean scores on the post-test also did not differ significantly between those with a pre-test and those without ($p=0.74$).

Missing Post-test

There was no difference in the mean age of those who took a post-test and those who did not, but participants under the age of 15 were 4.8 times as likely to be missing a post test ($p=0.0097$) and those under 15.5 were 2.4 times more likely ($p=0.017$). There were no significant differences between those with a post-test and those without based on race/ethnicity or gender.

Those without a post-test attended significantly fewer sessions than those with a post-test; 9.4 compared to 6.2 ($p<0.0001$). In addition, this relationship was seen when excluding those who missed

both tests. Those who missed the pre-test but took the post-test attended an average of 7.8 sessions, compared to 9.6 for those who took both tests (p=0.006).

The mean scores on the pre-tests differed significantly between students who took the pre-test but not the post-test and students who took both tests (p=0.020). For those with both tests, the mean score on the pre-test was 62.2 out of 80, compared to 58.4 out of 80 for those who were missing a post-test. Those who took both tests gave correct answers to 15.2 questions out of 20, while those missing a post-test gave an average of 14.1 correct answers. However, the difference in number of correct answers was not significant (p=0.080).

4.5.2: Changes

The following analysis uses an intent-to-treat strategy, looking at all participants and not separating them out based on differential intervention exposure. Analysis according to intervention exposure is reported in the next section.

Percent Correct

There is a significant difference in the **average number of questions correct** between the pre-test and the post-test. The average number of questions correct on the pre-test was 14.7 (73.4%), compared to 15.6 (77.8%) correct on the post-test. This difference is significant (p=0.013).

There is no significant difference in the **number of questions answered correctly in any of the 9 composite items**. For each item, the average number of questions correctly answered increased from pre-test to post-test, but none of these differences were significant. The closest these differences came to significance was for the Pregnancy composite item (p=0.067).

Table 23: Change in percent of questions answered correctly between pre-test and post-test

Item	% of questions correct on pre-test	% of questions correct on post-test	Difference	p-value
Condoms	67.17%	68.82%	1.65%	0.6692
STIs	67.73%	71.61%	3.88%	0.2340
Drugs	75.76%	79.57%	3.81%	0.3461
Pregnancy	82.95%	89.25%	6.30%	0.0668
Attitudes	76.14%	80.29%	4.15%	0.1489
Behaviors	68.94%	76.34%	7.40%	0.1369
Knowledge	72.94%	77.57%	4.63%	0.1061
Self-efficacy	81.06%	82.80%	1.74%	0.6311
Intentions	53.79%	59.68%	5.89%	0.1981
All questions	73.4%	77.8%	4.4%	0.0130

The percentage of students answering questions correctly increased significantly from pre-test to post-test for only one question. The proportion of respondents answering q**uestion 1** correctly increased from 85.8% on the pre-test to 97.9% on the post-test (p=0.0007). No other changes were significant. The proportion of students responding correctly to each question increased between the pre-test and post-test, but no changes were significant except for question 1. Questions 14 (p=0.066), 2 (p=0.067), and 16 (p=0.069) came the closest.

Table 24: Change in percent of correct responses to questions between pre-test & post-test

Question	% correct on pre-test	% correct on post-test	Difference	Odds Ratio	p-value
1	85.8%	97.8%	12.0%	7.468	0.0007
2	73.7%	83.0%	9.3%	1.737	0.0669
3	54.1%	58.1%	4.0%	1.172	0.2813
4	79.7%	81.7%	2.0%	1.138	0.3569
5	82.0%	88.2%	6.2%	1.638	0.1387
6	85.0%	88.0%	3.0%	1.302	0.2607
7	47.7%	48.4%	0.7%	1.025	0.4637
8	69.5%	76.3%	6.8%	1.416	0.1309
9	78.9%	80.6%	1.7%	1.111	0.3813
10	68.4%	69.9%	1.5%	1.071	0.4091
11	66.4%	72.0%	5.6%	1.302	0.1879
12	78.2%	78.3%	0.1%	1.004	0.4976
13	85.6%	90.3%	4.7%	1.566	0.1505
14	75.0%	83.5%	8.5%	1.685	0.0656
15	82.6%	90.1%	7.5%	1.726	0.0892
16	68.9%	78.0%	9.1%	1.596	0.0688
17	85.6%	87.0%	1.4%	1.120	0.3919
18	74.8%	81.5%	6.7%	1.483	0.1210
19	78.8%	85.9%	7.1%	1.633	0.0910
20	32.6%	38.0%	5.4%	1.270	0.2012

Average Score

The **average score out of 80** on the pre-test was 60.5, compared to an average score of 62.8 on the post-test. This difference was not significant, though it was nearly so (p=0.057).

There was a significant increase in the **average score on the pregnancy composite**, with an increase from 6.68 out of 8 to 7.13 (p=0.036). All other composite items saw increases in scores, except for self-efficacy, which saw a small decrease. None of these changes were significant.

Table 25: Change in average scores on composite items between pre-test and post-test

Item	Points possible	Score on Pre-test	Score on Post-test	Difference	p-value
Condoms	12	9.08	9.14	0.06	0.8371
STIs	20	15.90	16.06	0.16	0.7062
Drugs	8	6.58	6.79	0.21	0.2618
Pregnancy	8	6.94	7.30	0.36	0.0366
Attitudes	24	19.39	19.97	0.58	0.1586
Behaviors	8	6.05	6.45	0.40	0.1284
Knowledge	28	23.24	23.72	0.48	0.3000
Self-efficacy	12	9.98	9.95	-0.04	0.8764
Intentions	8	5.38	5.70	0.33	0.1420
All questions	80	60.50	62.80	2.30	0.0573

Average scores on all but two questions increased, most of them slightly. Two questions – question 7 and question 12 – saw very slight decreases. All of these changes were insignificant except one. **Average scores on question 1**, which asked whether a girl can get pregnant the first time she has sex, increased from 3.537 out of 4 to 3.935 (p=0.0006). Question 14 came the closest to having a significant change in score (p=0.056).

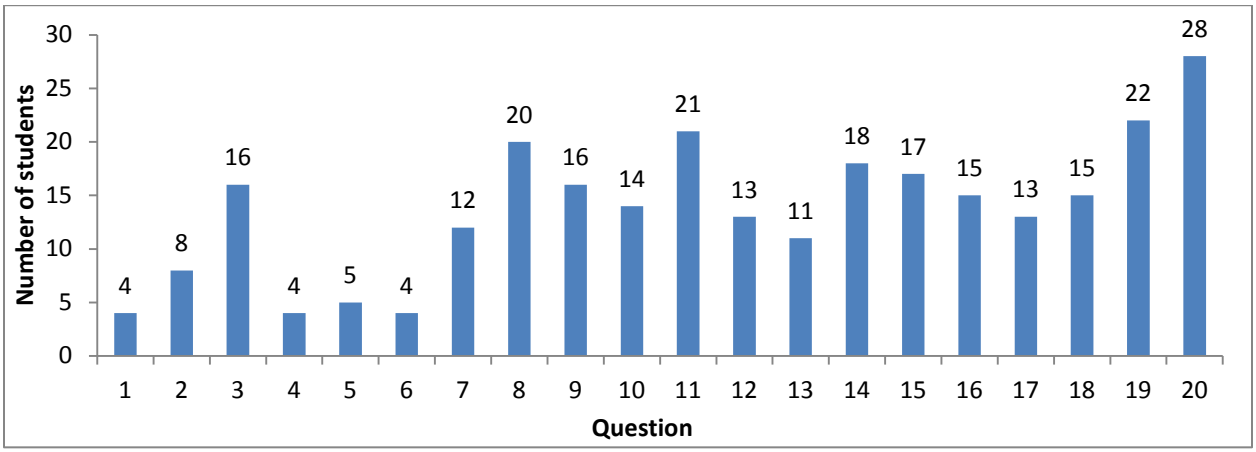
Table 26: Change in average scores on questions between pre-test and post-test

Question	Score on Pre-test	Score on Post-test	Difference	p-value
1	3.5373	3.9355	0.3982	0.0006
2	3.3985	3.5745	0.176	0.2144
3	2.5414	2.5806	0.0392	0.8659
4	3.5639	3.5699	0.006	0.9619
5	3.4436	3.6129	0.1693	0.2989
6	3.6241	3.6957	0.0716	0.5682
7	2.1667	2.1538	-0.0129	0.9592
8	2.771	2.9785	0.2075	0.2826
9	3.1504	3.1935	0.0431	0.7912
10	2.7519	2.7527	0.0008	0.9954
11	2.6641	2.8495	0.1854	0.3195
12	3.1805	3.1522	-0.0283	0.8514
13	3.2955	3.3763	0.0808	0.4968
14	3.0682	3.3297	0.2615	0.0556
15	3.3106	3.3516	0.041	0.7403
16	2.8788	3.1758	0.297	0.1163
17	3.3106	3.3152	0.0046	0.9789
18	3.0763	3.2174	0.1411	0.3965
19	3.053	3.0978	0.0448	0.7677
20	1.7045	1.9239	0.2194	0.2559

Positive Change

On average, participants **showed positive change** – that is, they moved closer to the desired answer between pre- and post-test – on 18.5% of questions. On 16.2% of questions, participants moved in a negative direction, and in the remainder, they did not move at all. The questions most likely to see change in a positive direction were question 20 (28 of 74), question 19 (22 of 74) and question 11 (21 of 74). The questions least likely to change included questions 1, 4, and 6; for all three, only 4 of 74 participants moved in a positive direction. All three of these questions, and most others as well, saw very little movement in a negative direction either. For these questions, 69, 64, and 65 saw no movement at all, respectively. The questions with the greatest movement in a negative direction included question 10 (21 of 74), question 12 (20 of 74), and question 19 (18 of 74).

Figure 11: Number of students showing positive change between pre-test and post-test



The likelihood of positive change was mostly unrelated to demographic factors. Age and race had no effect on a question showing positive change or not. However, differences in gender were seen for question 14 ("I would insist that my partner or I use a condom every time we had sex"); 36.7% of male students showed positive change for this question in comparison to 16.7% of female students (p=0.0494).

Distribution of Answer Choices

Four questions saw a **significant increase in the proportion of participants who achieved the "maximum" possible answer**. For questions 1 through 7, this figure is identical to the proportion who answered the question correctly; as seen above, questions 1 and 2 both saw significant increases in the proportion of respondents answering correctly. **Question 8** saw an increase in the proportion of participants responding "strongly agree," increasing from 44.3% on the pre-test to 53.8% on the post-test (p=0.083). **Question 16** saw an increase in the proportion of participants responding "strongly disagree," increasing from 52.3% in the pre-test to 65.9% in the post-test (p=0.0218).

Question 14 also saw a significant decrease in the proportion of participants responding either "strongly disagree" or "disagree." This differs from the ordinary measure of correct vs. incorrect

because it excludes those who responded “don’t know.” The proportion of participants responding as such decreased from 13.6% in the pre-test to 4.4% in the post-test.

Section 4.6: Effect of Intervention Exposure

Intervention exposure was quantified by classifying sessions as either “successful” or “unsuccessful.” The criteria used for this classification is discussed above in the methods chapter. In this section, we discuss the effect of attendance at successful sessions on various metrics. These include the numeric change in score – that is, the score on an item on the pre-test subtracted from the score on the post-test. This change was also dichotomized by characterizing it as “positive” or “not positive.” We also examine the relationship between successful session attendance and the proportion of items correct on the post-test, as well as the mean score on items in the post-test. In these two-subsections, stratified analysis was conducted to examine the effect of intervention exposure on post-test scores among those with high and low scores on the pre-test.

Each metric is also analyzed for the effect of simple attendance at sessions, rather than attendance at successful sessions.

4.6.1: Numeric Change in Score

Attending 8 or more successful sessions was significantly associated with more **desirable change in scores for question 15**. The mean change in scores for those attending 7 or fewer successful sessions was -0.31, compared to +0.23 for those attending 8 successful sessions or more (p=0.031). The same was true for the **change in score for question 4**. Those attending 7 or fewer successful sessions had a mean change in score of -0.33, compared to +0.091 for those attending 8 or more successful sessions (p=0.032).

The same pattern was seen in the **change in score on the STI composite item**. Those attending 7 or fewer successful sessions saw a change of +0.100, compared to +0.1867 for those who attended 8 or more successful sessions (p=0.040).

Simple attendance at sessions, whether successful or not, had no effect on the change in score between pre-test and post-test for question 15, question 4, or the STI composite.

4.6.2: Positive Change

There was no significant relationship between the proportion of questions seeing positive change and successful session attendance. Those attending 7 or fewer successful sessions saw positive change on 15.6% of questions on average, compared to 20.5% for those attending 8 or more successful sessions. However, this association was not significant (p=0.123). The same was true for the proportion of questions seeing negative change (p=0.442).

Among those answering **question 7** incorrectly in the pre-test, higher attendance of successful sessions strongly affected **positive change in the question score** from pre-test to post-test. Only 11.8% of those attending 7 or fewer successful sessions saw positive change in their question score, compared to 45.5% of those attending at least 8 sessions. (p=0.026). The same was seen for **question 9**; 100% of those attending 8 or more successful sessions saw positive change, compared to only 30% of those attending 7 or fewer (0=0.01).

Simple attendance at sessions, whether successful or not, had no effect on the change in score between pre-test and post-test for question 7 or question 9.

4.6.3: Correct Responses on Post-Test

Among those who answered less than 75% of the questions on the pre-test correctly, attending 8 or more sessions was strongly and significantly associated with **answering more than 75% of questions correctly on the post-test.** Only 30% of those who attended 7 or fewer successful sessions answered

more than 75% of questions on the post-test correctly, compared to 78.6% of those who attended 8 or more successful sessions (p=0.024). This relationship was not significant among those who answered more than 75% of pre-test questions correctly. Attending more sessions (whether successful or not) had no effect on the proportion of students answering more than 75% of questions correctly on the post-test.

There is a strong significant relationship between the number of successful sessions attended and the likelihood of **scoring correctly on question 7 on the post-test.** Among those who attended 8 or more successful sessions, 58.2% answered this question correctly, compared to only 33.3% of those who attended 7 sessions or fewer (p=0.017). This effect was even stronger for those who answered question 7 incorrectly on the pre-test. Among these students, those attending 8 or more successful sessions were 9.1 times more likely to answer the question correctly on the post-test, compared to those who attended 7 or fewer successful sessions (p=0.028). The same relationship did not hold for those who answered the question correctly on the pre-test.

Attending 8 or more sessions, whether the sessions were successful or not, was associated with answering question 7 correctly on the post-test. Among those who attended 7 or fewer sessions, 27.3% answered the question correctly on the post-test, compared to 55.1% of those attending 8 or more sessions (p=0.020). However, this relationship was not seen once the data was broken down according to correct answers on the pre-test. Neither those who answered question 7 correctly on the pre-test nor those who answered it incorrectly saw any significant relationship between attending 8 or more sessions and answering the question correctly on the post-test (p=0.194, p=0.295).

A similar effect was seen on **question 15**. Among those who attended 8 or more successful sessions, 98.1% answered this question correctly, while among those that attended 7 successful sessions or fewer, 78.4% answered this question correctly (p=-0.003). This relationship was seen both in those that

answered the question correctly on the pre-test and those that did not, though small sample sizes made stratified analysis impossible for this question.

Attending 8 or more sessions (whether successful or not) was associated with a higher proportion of students answering question 15 correctly on the post-test (p=0.004). Only 71.4% of those attending 7 or fewer sessions answered the question correctly on the post-test, compared to 95.7% of those attending 8 or more. This relationship did not hold up for those who answered the question incorrectly on the pre-test (p=0.491). However, the association did hold up for those who answered the question correctly on the pre-test. The proportions were similar, with 75% of those attending 7 or fewer sessions answering correctly on the post-test, compared to 98% of those attending 8 or more sessions (p=0.021).

Attending more than 7 successful sessions was also a significant predictor of **correctly answering at least 75% of questions within the STI composite item correctly on the post-test**; 44.7% who attended 7 sessions or less answered at least 75% correctly, compared to 65.5% of those attending 8 successful sessions or more (p=0.038). Interestingly, this relationship did not hold up for those who answered less than 75% of questions correctly on the pre-test; in this group, attending 8 or more successful sessions was associated with a 2.75 times greater likelihood of answering more than 75% of questions correctly on the post-test, but this association was not significant (p=0.173). However, among those who did answer more than 75% of questions correctly on the pre-test, attendance at successful sessions was strongly associated with maintaining a high score on the post-test (OR=10.5, p=0.006). Among those with more than 75% correct on the pre-test, those who also achieved this score on the post-test attended an average of 9.3 successful sessions, much higher than the average of 4.6 attended by those whose post-test score dropped below 75% of questions correct (p=0.0003).

Attending 8 or more sessions (whether successful or not) was not significantly associated with correctly answering more than 75% of questions within the STI composite item on the post-test(p=0.068), but this

relationship was significant among those answering more than 75% correctly on the pre-test (p=0.007). Only 28.6% of those attending 7 or fewer sessions answered 75% or more of questions correctly, compared to 84.4% of those attending 8 or more sessions. The relationship was insignificant among those answering less than 75% of questions correctly on the pre-test (p=0.571).

The same was true with questions within the **knowledge composite item on the post-test**; 42.1% of those attending 7 or fewer successful sessions answered at least 75% of questions correctly, compared to 63.6% of those attending 8 or more successful sessions (p=0.033). This item saw the same pattern as seen with the STI composite item above. The relationship did not hold up for those answering less than 75% of questions correctly on the pre-test (p=0.219). Those answering more than 75% of questions correctly on the pre-test were 8.9 times more likely to do so on the post-test if they attended 8 or more successful sessions than if they attended 7 or fewer (p=0.007). Attendance at more successful sessions also helped those answering more than 75% of questions correctly on the pre-test maintain this score. Within this group, those that dropped below 75% on the post-test attended an average of 5.2 successful sessions, compared to 9.6 for those who maintained a score of more than 75% of questions correct.

There was no relationship overall between attendance at sessions (whether successful or not) and answering more than 75% of questions correctly within the knowledge composite item on the post-test (p=0.104). Again, this relationship was not significant among those answering less than 75% of questions dealing with knowledge correctly on the pre-test (p=0.592), but was significant for those answering more than 75% of pre-test knowledge questions correctly (p=0.016). Within this group, 28.6% of those attending 7 or fewer sessions answered more than 75% of questions correctly, compared to 80% of those attending 8 or more sessions.

4.6.4: Average Scores

Among those who answered less than 75% of questions on the pre-test correctly, the **mean score on the post-test** was significantly higher for those attending at least 8 successful sessions; 62.14 out of 80 compared to 54.69 for those attending 7 or fewer (p=0.039). This relationship did not hold for those answering more than 75% of questions correctly on the pre-test (p=0.380) or overall (p=0.177). Attendance at sessions (whether successful or not) had no effect on the mean score on the post-test, whether overall or divided based on score on the pre-test.

The **average score on the behavior composite item on the post-test** varied strongly according to number of successful sessions attended. Those attending 7 or fewer successful sessions had an average score of 7.4 points out of 10, compared to 8.5 points for those attending 8 or more sessions (p=0.019). **Among participants who answered less than 75% of these questions correctly in the pre-test**, the difference was even more striking. Those attending 7 or fewer successful sessions had an average score of 5.8 out of 10, compared to 8.2 for those attending 8 or more successful sessions (p=0.015).

Attendance at sessions (whether successful or not) had a significant effect on the average scores on the behavior composite item on the post-test, both for those answering less than 75% of questions correctly on the pre-test, those answering more than 75% correctly, and overall. Overall, the mean score for those attending 7 or fewer sessions was 6.8 out of 10, compared to 8.5 for those attending 8 or more sessions (p=0.003). A similar pattern was seen for those answering more than 75% of behavior items correctly on the pre-test; the mean score for those attending 7 or fewer sessions that 7.8 compared to 9.1 for those attending 8 or more (p=0.034). Among those answering less than 75% of behavior items correctly on the pre-test, the average score on the behavior item on the post-test was 5.3 out of 10 for those attending 7 or fewer sessions, compared to 7.7 for those attending 8 or more (p=0.047).

Those attending 7 or fewer successful sessions had an **average score** of 19 points out of 25 on **the STI composite item on the post-test**, compared to 20.8 points for those attending 8 or more successful sessions (p=0.008). **This relationship was seen for both** those answering less than 75% correct on the pre-test (p=0.019) and those scoring more (p=0.020). Attendance at sessions (whether successful or not) had no effect on the mean score for items dealing with STIs, whether overall or when divided by score on the pre-test.

Attendance at successful sessions had a significant effect on the **average score on the knowledge composite item on the post-test**. Those attending 8 or more successful sessions had an average score of 30.6 out of 35, compared to 28.3 for those attending 7 or fewer sessions (p=0.006). Attendance at sessions (whether successful or not) had no effect on the mean score for items dealing with knowledge, whether overall or when divided by score on the pre-test.

Among those answering question 9 incorrectly in the pre-test, attending 8 or more successful sessions was strongly associated with **a higher score on this item in the post-test;** those attending 8 or more successful sessions had a mean score of 4.5, compared to a mean score of 2.8 for those attending 7 or fewer successful sessions (p=0.033). The same association was not seen among those who answered the question correctly in the pre-test. The mean scores on question 9 in the post-test were not significantly affected by attendance at sessions (whether successful or not).

Chapter 5: Discussion

Section 5.1: Demographic Characteristics

The average age and age distribution of the O2T participants under study is significantly higher than the intended age the program was designed for. The program description prescribes a possible range of 9 to 18 years-old. In addition, many of the activities are geared towards a younger audience. In session 1, activity 1.4 “My Future” asks participants to imagine themselves in high school in the future. Session 2 also spends a significant amount of time discussing puberty, a topic that is usually covered in middle school or even elementary school. Anecdotally, facilitators reported that many classes lost interest in the sections of the curriculum that discusses puberty because they had already learned this material, or having passed through the majority of physical changes already, felt the information to be irrelevant to them. The average age of O2T participants is 16.5. Only 15 of 178 students (8.4%) were under the age of 15, and the oldest student was nearly 21. It is expected that the average age of participants would be high; four of the eleven series, with a total of 101 participants, are located in classes in John Harris High School. Furthermore, because of the low four-year graduation rate, many students at John Harris are older than is typical. Of this 96 students with ages recorded in the John Harris cohort, 23 (24%) are 19 or over. This age difference may have been one of the reasons why the curriculum was less effective than hoped for.

More than half of participants were missing information on their race or ethnicity. For those with race recorded, the distribution is somewhat representative of the areas in which the programs were held. Harrisburg city is 16.2% Hispanic and 50% non-Hispanic Black. Steelton borough is 11.8% Hispanic and 41.2% non-Hispanic Black. Susquehanna Township is 3.4% Hispanic and 21.3% non-Hispanic Black. Muncy borough is 96.7% White, but residents of Ashler Manor originate from all over Lycoming County and from other counties as well. If anything, the race/ethnicity composition of program participants over-represented

Section 5.2: Intervention Exposure

The intervention was conducted with less fidelity than is ideal. Six activities were omitted in more than half of all series. Only 17 activities were conducted as described in a majority of series.

Activity 2.4 was omitted in all but one series. This activity involves a description of the biology of vaginal intercourse and the process of fertilization, including a short video. In all of the New Beginnings series, facilitators were not permitted to conduct this activity at all. At one of the John Harris series, disruptive behavior by students interrupted the session and made completion of this activity impossible. At the other John Harris series, the school requested that school board-approved material be used instead of the O2T video. Only at Ashler Manor was this activity conducted as described.

Activity 3.1 involved a demonstration of how to put on a condom and a memory game using the steps in doing so. The use of an actual condom was not permitted at any of the New Beginnings sessions, and in other series the activity was modified. This activity was only conducted as described once; it was omitted five times and changed five times.

Other activities were implemented with low fidelity due to unenthusiastic student participation or high absentee rates making certain exercises difficult. Activity 5.2 involves the facilitator answering questions submitted by students about drugs and alcohol. Students often did not participate in this activity; they did not submit questions ahead of time and did not volunteer questions when given the opportunity to do so. Three activities in session 6 involve role-play and group discussion. If attendance at sessions was very low or students were not interested or willing to participate, these activities were often either omitted or altered significantly.

Nearly a third of recorded students attended 4 sessions or fewer, including about 6% who attended no sessions at all. Most of these latter students are composed of those who registered for the class and so were on the records, but withdrew from the class before it began. Some of these students were

expected to return to class at some point following out-of-school suspension or maternity leave, and so were kept on the rosters despite never attending a class. It is also possible that mistakes were made in recording attendance; these students may have in fact attended at least one session, but were not recorded properly and thus the data reflect them as attending no sessions. On average, about 58% of enrolled students were present at each session. This ranges from about 45% to 74% across sites. However, the high absentee rate does not in itself result in unsuccessful sessions; while absent students are not exposed to that session's content, as long as at least three or four students are in attendance those present can benefit from the session. The high absentee rate causes problems and leads to failed sessions in conjunction with small numbers of enrolled students. One of the Fall 2013 John Harris series had only 6 students enrolled; if about half of students are absent at any given session, this increases the odds that attendance will drop below feasible levels. The problem is compounded when multiple sessions are consolidated into one day – if that happens to be a day with many absences, many students miss out on the material covered and even those present do not benefit as much. For example, in the Summer 2012 Steelton New Beginnings series, sessions 4, 5, and 6 were compressed into one day. While consolidating material like this is a problem in and of itself, it became particularly problematic when only one of the five enrolled students was present that day and refused to participate in most of the activities.

Section 5.3: Test Scores

From an intent-to-treat perspective, a few changes from pre-test to post-test were significant. The average number of questions correct increased significantly, as did the proportion of students correctly answering questions 1 and 2. These questions are, respectively, "Can a girl get pregnant the first time she had sex?" and "Can you get STIs from having oral sex?" The average score on question 1 also increased significantly. For the "pregnancy" index, which incorporated both question 1 and question 9, the average score also increased. Question 9 asks respondents to agree or disagree with the statement

"Getting pregnant or getting someone pregnant at my age would make it more difficult to achieve my goals."

Using an intent-to-treat analysis, comparing pre-test and post-test scores suggests that the intervention was successful in improving knowledge on two specific facts about STI acquisition and pregnancy, and potentially improving attitudes surrounding pregnancy and future goals.

Two additional questions saw significant increases in the proportion of students who responded with the "most correct" answer. Both of these questions assessed behavior; between pre-test and post-test, the number of students who "strongly disagreed" with the statements "In the past month I have had sex" (question 8) and "During the past month I drank alcohol, used marijuana, cocaine, or street drugs" (question 16) increased significantly.

While the intervention did not appear to decrease the proportion of students reporting sexual behavior or drug use, it seems to have, at the least, increased the vehemence of their denial of such behavior.

It is unclear why this question was chosen to be answered on a strongly agree – strongly disagree spectrum rather than a yes/no or true/false. Indeed, what does it mean if a student "disagrees" with the statement that they were sexually active in the last month? How does this differ from "strongly disagreeing"? Does one imply a greater or lesser level of sexual activity? Anytime study participants are asked to self-report potentially risky or stigmatizing behaviors, the possibility for social desirability bias exists. Students may report that they have not used drugs, even if they have, if they believe this is what the evaluator wants to hear. On the other hand, social pressures may lead adolescents to over-report certain activities. In examining the breakdown of responses across all 13 Likert scale variables, it is interesting to observe which questions have the largest number of students responding with the "most incorrect" answer. Question 8, as stated above, asks respondents to agree or disagree with the statement that they have been sexually active in the last month. More than 12% of students responded

"strongly agree." Question 20 asks respondents to agree or disagree with the statement "I plan to wait until I am older to have sex." More than 24% of students "strongly disagreed" with this statement. These proportions are the highest of any question, reflecting larger numbers of students choosing the "most incorrect" answer. Responses to this question, of course, are very sensitive to the student's age. A 13-year-old "strongly disagreeing" with this statement has a very different meaning than a 19-year-old doing so.

The three questions with both the lowest average scores and the lowest percentage of students answering correctly are questions 3, 7, and 20. Question 3 asks students "Can you get HIV from kissing someone?" Question 7 is also a "knowledge" question, asking "Can wearing a condom keep you from getting an STI or HIV?" It is interesting and troubling that students scored so low on this question. Question 3 may be poorly designed. The "correct" answer is actually slightly unclear, as transmission of HIV by kissing is theoretically possible if both parties have open sores in their mouths but has never once been documented (it is also unclear if the question assumes mouth-to-mouth kissing). Arguably the distinction between activities that transmit HIV and those that do not could have been more relevantly drawn elsewhere. However, question 7 is far more straightforward and is arguably a far more important piece of knowledge for engaging in safe sexual activity. Less than 50% of students answered this question with "Yes"; while about 13% were unsure, nearly 40% answered "No."

Intervention exposure had significant effects on both the percent of questions answered correctly and the mean score on the post-test. Those who attended 8 or more successful sessions had significantly higher scores and higher percentages of questions correct on the post-test than those who attended fewer successful sessions. What is very interesting about this relationship is that it both cases it only holds for those answering less than 75% of questions correctly on the pre-test.

This shows that intervention exposure is associated with significant improvements among those who entered the program with low scores on the pre-test. Across several variables, analysis shows that the same effects were not seen for those entering with high scores. This is an important and positive finding, as it shows that the intervention was able to benefit those who most needed it.

Most of the items that showed improvement in response to intervention exposure dealt with knowledge. Improvements in several measures were seen for questions 4 ("Can a person have an STI and not know it?"), question 7 ("Can wearing a condom keep you from getting an STI or HIV?") and question 15 ("If a person has unprotected sex one time, s/he can get an STI or HIV"). Improvements were also seen in the index scores that measure knowledge about STIs and that measure all knowledge items, including those dealing with STIs, condoms, pregnancy, and drugs.

Two important exceptions to this were improvements in question 9 ("Getting pregnant or getting someone pregnant at my age would make it more difficult to achieve my goals") and the index item measuring behavior. This index included self-reported drug or alcohol use (question 16) and sexual activity (question 8) in the past month. These items asked students to agree or disagree with the statements "In the past month I have had sex (vaginal, oral, or anal)" and "During the past month I drank alcohol, used marijuana, cocaine, or street drugs." Both of these items improved significantly in response to intervention exposure. Among those who answered question 8 "incorrectly" on the pre-test – that is, those who responded the statement with "strongly agree", "agree", or "don't know" – only 40% of those attending 7 or fewer successful sessions "disagreed" or "strongly disagreed" with the statement on the post-test compared to 62.5% of those who attended 8 or more successful sessions. Similarly, among those answering question 16 "incorrectly" on the pre-test, 33% of those attending 7 or fewer sessions "disagreed" or "strongly disagreed" with the statement on the post-test compared to 80% of those attending 8 or more sessions.

It is common for teen pregnancy programs to show improvement only in items dealing with knowledge but not to show any significant changes in behavior. This finding shows that the intervention did result in significant and important changes in reported risky behavior.

An important finding is that for most cases in which improvement was seen in response to attendance at successful sessions, no improvement was seen for the simpler metric of attendance at sessions, whether those sessions were successful or unsuccessful. This shows that the quality of each session is nearly as important as whether an individual student attends that session. This depends both on the facilitator and on other students; fidelity to curriculum, session attendance, and student participation and interest are all critical to ensuring that students receive benefit from the program.

It is important to note that for almost all metrics, improvement was seen between pre-test and post-test. However, in the vast majority of cases, these changes were not significant. This is greatly due to the poor quality of the data. Only 213 students were recorded in total, and many of these missed the pre-test, post-test, or both. A mere 74 students took both tests, resulting in a very small sample size for assessing change between pre-test and post-test.

Chapter 6: Recommendations and Conclusion

Section 6.1: Recommendations for Program Implementation

This evaluation found that this program was often implemented with low fidelity to curriculum. On average, less than half of activities were conducted as described in the curriculum, and about 21% were never conducted at all. We recommend that FHCCP re-evaluate the curriculum in partnership with facilitators, and discuss why certain activities are so frequently changed or omitted. The results of this evaluation show that high fidelity to curriculum is a necessary component of the desirable changes seen in response to the intervention.

Another pernicious issue is low attendance at sessions, and lack of participation by students who are present. These issues go hand-in-hand, as low participation and levels of interest were often seen in conjunction with low attendance. Many of the activities in the curriculum are impossible to conduct without at least 3 or 4 students present, and groups discussion activities benefit strongly from larger groups. FHCCP should examine, in partnership with facilitators and sites, how to increase attendance at each session. It may be valuable to seek feedback from students explaining why they happen to miss sessions. Lessons have already been learned from the first year of New Beginnings series that facilitators and program staff must work closely with site staff to ensure complete understanding of the program and to set a firm schedule for each session. Especially with New Beginnings, sessions were often missed because site staff did not communicate with the facilitator about other activities students would be involved in, special events, and days when the camp was not in session. This is a particular problem when the last session is missed, as happened with one New Beginnings series in 2013, because this means the loss of not only instructional time but post-test data.

In addition to ensuring effective communication regarding session schedules, FHCCP should work more closely with sites to ensure understanding of program content. In the New Beginnings series, facilitators

were prohibited from conducting certain activities on birth control and pregnancy because their content was perceived as inappropriate by New Beginnings Staff. In one case, site staff were not opposed to curriculum content in general, but prohibited conducting certain activities because the O2T sessions were conducted in the same room as younger children.

Along with the issue of participation, facilitators sometimes encountered problems with disruptive student behavior, particularly in response to certain activities. In one of the John Harris series, disruptive behavior prevented a number of activities from being conducted during one session. At the beginning of the subsequent session, the regular class teacher (a John Harris employee) spoke to the students and ensured much better behavior during that session. This anecdote shows that site staff have an important role to play in controlling disruptive behavior and improving the implementation of the program. It may also be appropriate to include classroom control techniques in facilitator training.

Section 6.2: Recommendations for Further Evaluation

At the moment, the pre- and post-tests for this program are proctored by the same facilitator who conducts the educational activities throughout the 7 sessions. To avoid social desirability bias, it would be better to have another person administer the tests. At a minimum, this should be someone other than the main facilitator for the group; ideally, this will be someone completely external to the project, with whom the participants will interact only once.

Pre- and post-test administration is hampered by the number of students missing one or both tests, despite having participated in most of the sessions; students that miss the first or last sessions do not have an opportunity to “make up” the test at another time. We recommend that students take the pre-test at the first session they attend, even if this is not the first session for that particular series. Post-tests were also sometimes omitted because students missed the last session or the entire session was cancelled. This may bias the data, as those who are more likely to attend these sessions may also be pre-

disposed to improvement in scores (or vice versa). For example, if a student misses both the first and last sessions but attends at least some in-between, he will be effectively invisible to the evaluation. If he misses only the last session, the potential improvement in his scores will not be measured. If a student is more likely to miss the first session and also more likely to have a lower pre-test score, the overall pre-test scores will overestimate existing knowledge/attitudes.

Better coordination of schedules and timing should help address the latter problem. Students should also be given the post-test if they know ahead of time that they will miss the last session. This may cause difficulties when considered in conjunction with the previous recommendation: an external test proctor may not be available at sessions other than the first and last. In this case, the facilitator must make their best effort to have the test administered by an appropriate person.

However, increasing pre- and post-test completion may be useless if the differences in proctoring introduce biases; some tests will be proctored by an outside person and others by the main facilitator or another staff member. A better strategy may be to have all tests proctored by the facilitator or the same site staff member, but reduce bias by ensuring student confidentiality. Confidentiality measures, such as de-identification of data, should be explained fully to the students. Other measures include using a cover sheet with the survey, so that answers are not visible to others.

In addition, information on race/ethnicity and date of birth is often not recorded for students who miss the first session. Even if ensuring pre-test completion by those missing the first session is not possible, at a minimum, this information should be collected as soon as possible to improve the completeness of data.

Variability both in post-test scores and in the change between pre- and post-test may be due to external factors that also affect intervention exposure. For example, students may be both more likely to miss sessions and less likely to show improvement in test scores over time due to some other characteristic.

Another issue hinges on the fact that currently, missing the first or last session means that a student does not take the pre- or post-test. This may bias the data, as those who are more likely to attend these sessions may also be pre-disposed to improvement in scores (or vice versa). For example, if a student misses both the first and last sessions but attends at least some in-between, he will be effectively invisible to the evaluation. If he misses only the last session, the potential improvement in his scores will not be measured. If a student is more likely to miss the first session and also more likely to have a lower pre-test score, the overall pre-test scores will overestimate existing knowledge/attitudes.

This curriculum is designed specifically for young teens; however, in practice the ages of the participants have ranged from 12 to 19, with an average age between 14 and 15. The results of this evaluation therefore may not be generalizable to the intended age group. It was initially geared towards an urban setting with mostly Black and Hispanic students, but as part of this project is being implemented in more rural and suburban settings as well. It may be feasible to look at these general populations separately, and determine, for example, that the intervention is effective among Black students in an urban setting but not otherwise.

The evaluation plan already in place does not provide for follow-up at a later point. This is a limitation, as most other studies include at least one follow-up and usually several, from 6 weeks post-intervention to 4 years. Another limitation is the inconsistent timing of the intervention. Most series are delivered in about 7 weeks, with one session each week. However, some have been delivered in as a little as two weeks or over up to 4 months. Temporal distance from previous sessions at post-test may introduce biases.

Section 6.3: Conclusion

The value of this evaluation is limited by the poor quality of the data, resulting in few significant findings. However, despite the small effective sample size and quantity of missing data, this evaluation found

several strong and significant effects of the intervention. Significant effects were most often seen for metrics measuring knowledge; intervention exposure led to increased knowledge about STIs, condoms, and pregnancy. However, intervention exposure was also associated with desirable changes in attitudes surrounding adolescent pregnancy and future goals, and with reduced reporting of drug or alcohol use and sexual activity within the prior 30 days. This finding shows that O2T may overcome some of the limitations typically found in short-term, classroom-based pregnancy prevention programs, which typically struggle to show significant changes in behaviors. Overall, however, the findings from this evaluation are consistent with the literature on this topic, and show that O2T is a promising intervention that deserves further attention.

Bibliography

1. *Dropouts Public by School, 2011-2012*, P.D.o. Education, Editor 2013.
2. *Actual Four-year Graduation Rate, 2011-2012*, P.D.o. Education, Editor 2013.
3. *Policy Brief: Preventing Teen Pregnancy is Critical to School Completion*, 2010, The National Campaign to Prevent Teen and Unplanned Pregnancy.
4. *County Health Profiles*, P.D.o.H.B.o.H.S.a. Research, Editor 2012: Harrisburg, PA.
5. *National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention: ATLAS*. 2013; Available from: http://gis.cdc.gov/GRASP/NCHHSTPAtlas/main.html.
6. *American Community Survey 3-year Demographic and Housing Estimates & Economic Estimates: Harrisburg City School District, Dauphin County, PA.*, U.C. Bureau, Editor 2011.
7. *Teen Sexual Behavior and Contraceptive Use: Data from the Youth Risk Behavior Survey, 2011*, in *Fast Facts*2012, The National Campaign to Prevent Teen and Unplanned Pregnancy.
8. Inman, D.D., et al., *Evidence-Based Health Promotion Programs for Schools and Communities.* American Journal of Preventive Medicine, 2011. **40**(2): p. 207-219.
9. Jones, J., Mosher, W., Daniels, K. , *Current contraceptive use in the United States, 2006-2010, and changes in patterns of use since 1995*, in *National Health Statistics Reports*2012, National Center for Health Statistics: Hyattsville, MD.
10. Martin, J.A., Hamilton, B. E., Osterman, M. J. K., Curtin, S. C., Mathews, T. J., *Births: Final Data for 2012*, in *National Vital Statistics Reports*2013, National Center for Health Statistics, Centers for Disease Control and Prevention.
11. Mosher, W.D., Jones, J., Abma, J. C.. , *Intended and Unintended Births in the United States: 1982-2012*, in *National Health Statistics Reports*2012, National Center for Health Statistics: Hyattsville, MD.
12. Curtin, S.C., Abma, J. C., Ventura, S. J., Henshaw, S. K., *Pregnancy Rates for US Women Continue to Drop*, in *National Center for Health Statistics Data Briefs*2013, National Center for Health Statistics, Centers for Disease Control and Prevention: Hyattsville, MD.
13. Kost, K., Henshaw, S., *U.S. Teenage Pregnancies, Births and Abortions, 2008: State Trends by Race and Ethnicity*, 2013, Guttmacher Institute: Washington, D.C.
14. Sieving, R.E., et al., *Prime Time: 12-Month Sexual Health Outcomes of a Clinic-based Intervention to Prevent Pregnancy Risk Behaviors.* Journal of Adolescent Health, 2011. **49**(2): p. 172-179.
15. *Counting it up: The Public costs of teen childbearing: key data*, 2013, The National Campaign to Prevent Teen and Unplanned Pregnancy: Washington, DC.
16. DiClemente, R.J., L.F. Salazar, and R.A. Crosby, *A Review of STD/HIV Preventive Interventions for Adolescents: Sustaining Effects Using an Ecological Approach.* Journal of Pediatric Psychology, 2007. **32**(8): p. 888-906.
17. *Sexually Transmitted Disease Morbidity for selected STDs by age, race/ethnicity and gender 1996-2011*, 2012, US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for HIV, STD and TB Prevention (NCHSTP), Division of STD/HIV Prevention.
18. *Incidence, Prevalence, and Cost of Sexually Transmitted Infections in the United States*, in *CDC Fact Sheet*2013, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention.
19. Jim Braxton, e.a., *Sexually Transmitted Disease Surveillance 2012*, 2014, Centers for Disease Control and Prevention: Atlanta, GA.
20. Gavin, L.E., et al., *A Review of Positive Youth Development Programs That Promote Adolescent Sexual and Reproductive Health.* Journal of Adolescent Health, 2010. **46**(3, Supplement): p. S75-S91.

21. *Youth Risk Behavior Surveillance - United States, 2011*, in *Morbidity and Mortality Weekly Report*2012, Centers for Disease Control and Prevention: Atlanta, GA.
22. Sabia, J.J., *Does sex education affect adolescent sexual behaviors and health?* Journal of Policy Analysis and Management, 2006. **25**(4): p. 783-802.
23. Martinez, G., Abma, J., Copen, C., *Educating teenagers about sex in the United States*, in *NCHS Data Briefs*2010, National Center for Health Statistics: Hyattville, MD.
24. Zimmerman, R.S., et al., *Effects of a School-Based, Theory-Driven HIV and Pregnancy Prevention Curriculum.* Perspectives on Sexual and Reproductive Health, 2008. **40**(1): p. 42-51.
25. Dailard, C., *Sex education: politicans, parents, teachers and teens.* The Guttmacher Report on Public Policy, 2011. **4**(1).
26. Bennett, S.E. and N.P. Assefi, *School-based teenage pregnancy prevention programs: A systematic review of randomized controlled trials.* Journal of Adolescent Health, 2005. **36**(1): p. 72-81.
27. Lindberg, L.D. and I. Maddow-Zimet, *Consequences of Sex Education on Teen and Young Adult Sexual Behaviors and Outcomes.* Journal of Adolescent Health, 2012. **51**(4): p. 332-338.
28. Cohen, S.A., *The Obama Administration's First Budget Proposal Prioritizes Sex Education and Family Planning but Not Abortion Access.* Guttmacher Policy Review, 2009. **12**(2).
29. *Sex and HIV Education*, in *State Policies in Brief*2014, Guttmacher Institute: Washington, D.C.
30. Finer, L.B. and J.M. Philbin, *Sexual Initiation, Contraceptive Use, and Pregnancy Among Young Adolescents.* Pediatrics, 2013.
31. Tebes, J.K., et al., *Impact of a positive youth development program in urban after-school settings on the prevention of adolescent substance use.* Journal of Adolescent Health, 2007. **41**(3): p. 239-247.
32. Gruchow, H.W. and R.K. Brown, *Evaluation of the Wise Guys Male Responsibility Curriculum: Participant-Control Comparisons.* Journal of School Health, 2011. **81**(3): p. 152-158.
33. Kirby, D., *Abstinence, Sex, and STD/HIV Education Programs for Teens: Their Impact on Sexual Behavior, Pregnancy, and Sexually Transmitted Disease.* Annual Review of Sex Research, 2007. **18**(1): p. 143-177.
34. *State and County Quick Facts: Harrisburg, PA*, U.C. Bureau, Editor 2013.
35. *American Community Survey 3-year Demographic and Housing Estimates: Dauphin County, PA.*, U.C. Bureau, Editor 2011.
36. Levy, F., *America's Best Places to Raise a Family*, in *Forbes*2010.
37. *Harrisburg area ranked among Top 10 recession-proof cities*, in *The Patriot News*2010.
38. Malawskey, N., *Harrisburg's eye-popping debt*, in *The Patriot News*2012.
39. Malanga, S., *The many ways that cities cook their bond books*, in *Wall Street Journal*2013.
40. Southwick, R., *Newsmakers 2012: Harrisburg and its financial woes continue to grow*, in *The Patriot News*2012.
41. *Guide to Pennsylvania Schools Statewide Ranking*, in *Pittsburgh Business Times*2013.
42. *Pennsylvania Department of Education: Academic Achievement Report 2011-2012, Harrisburg City SD*. 2013; Available from: http://paayp.emetric.net/District/Overview/c22/115222752.
43. Baker, B. *Stephen Reed, Mayor of Harrisburg*. 2006; Available from: http://www.citymayors.com/mayors/harrisburg_mayor.html.
44. *TPP Research and Demonstration Programs*. Office of Adolescent Health, US DHHS 2013; Available from: http://www.hhs.gov/ash/oah/oah-initiatives/teen_pregnancy/about/research-demonstration.html.
45. *Personal Responsibility Education Program Fact Sheet*, A.f.C.a.F. Family and Youth Services Bureau, US DHHS, Editor 2013.
46. *Web Page "Grants"*, U.D. Office of Minority Health, Editor 2012.

47. Adamchak, S., Bond, K., *"A Guide to Monitoring and Evaluating Adolescent Reproductive Health Programs."* F.o.Y.A. Pathfinder International, Editor 2000: Washington, DC.
48. Chinman, M., et al., *Enhancing Quality Interventions Promoting Healthy Sexuality (EQUIPS): A Novel Application of Translational Research Methods.* Clinical and Translational Science, 2013. **6**(3): p. 232-237.
49. Feinberg, M., et al., *Effects of the Communities That Care Model in Pennsylvania on Youth Risk and Problem Behaviors.* Prevention Science, 2007. **8**(4): p. 261-270.
50. Jemmott, J.B., *The Reasoned Action Approach in HIV Risk-Reduction Strategies for Adolescents.* The ANNALS of the American Academy of Political and Social Science, 2012. **640**(1): p. 150-172.
51. Boyce, W., et al., *Sexual Health of Canadian youth: Findings from the Canadian Youth, Sexual Health and HIV/AIDS Study.* Canadian Journal of Human Sexuality, 2006. **15**(2): p. 59-68.
52. Howard, M., et al., *Young males' sexual education and health services.* American journal of public health, 2004. **94**(8).

Appendices

Appendix 1: Fidelity Monitoring Forms

FIDELITY MONITORING

On Second Thought

(O2T)

Agency/ School District	
(Agency/School) Program	
Date (s)	
Educator	

Instructions

- Upon completion of each O2T session, please complete the corresponding fidelity monitoring form from this packet.
- You should have a packet for each group of students that you teach.
- Each form contains an activity grid, which provides an opportunity for you to give feedback on each activity within the session. For each activity, please indicate whether you taught the activity as described, taught the activity with changes, or did not teach the activity.
- Please describe in the "Notes" section for each activity your reasons for any eliminations, additions, or changes made. Comments and suggestions concerning the program content, structure, and clarity of the materials are particularly helpful. Also include any problems you faced (i.e., time constraints) and any suggestions for how to enhance the sessions.
- Be sure to complete the form promptly. Complete the form while you are teaching the session, or immediately following presenting the material so that your experiences are fresh in your mind.
- Provide as much feedback as possible. The more feedback you provide, the more helpful this evaluation tool will be in future implementations. You may attach additional pages to provide additional comments if necessary.

SESSION 1: On Second Thought (O2T) Introduction and Self Esteem

Date(s) Taught:		**Time Started:**	
Number of Students Present:		**Time Ended:**	

Were the participants interested?	**Did the students participate?**	**Did the session feel rushed?**
☐ Not at all interested ☐ Some were interested ☐ All or almost all were interested	☐ None or very few participated ☐ Some participated ☐ All or nearly all participated	☐ Very rushed ☐ Somewhat rushed ☐ Not rushed at all

Activity 1 **Business First:** **Welcome and Pre-test**	*Activity 2* **Our Rules**	*Activity 3* **What is *O2T*?**
☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach
NOTES:	NOTES:	NOTES:

Activity 4 **My Future**	*Activity 5* **Feeling Good About Me**	*Activity 6* **Session 1 Closure**
☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach
NOTES:	NOTES:	NOTES:

ADDITIONAL Session Notes:

SESSION 2: Puberty and Teen Pregnancy Prevention

Date(s) Taught:		**Time Started:**	
Number of Students Present:		**Time Ended:**	

Were the participants interested?	**Did the students participate?**	**Did the session feel rushed?**
☐ Not at all interested ☐ Some were interested ☐ All or almost all were interested	☐ None or very few participated ☐ Some participated ☐ All or nearly all participated	☐ Very rushed ☐ Somewhat rushed ☐ Not rushed at all

Activity 1 **Goal Review**	*Activity 2* **His and Hers:** **You Need to Know ALL Body Parts**	*Activity 3* **Going through Changes**
☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach
NOTES:	NOTES:	NOTES:

Activity 4 **Putting it All Together…** **How Babies are Made**	*Activity 5* **Why Wait? (To Become a Parent)**	*Activity 6* **Session 2 Closure**
☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach
NOTES:	NOTES:	NOTES:

ADDITIONAL Session Notes:

SESSION 3: Preventing Pregnancy and Promoting Healthy Relationships

Date(s) Taught:		**Time Started:**	
Number of Students Present:		**Time Ended:**	

Were the participants interested?	**Did the students participate?**	**Did the session feel rushed?**
☐ Not at all interested ☐ Some were interested ☐ All or almost all were interested	☐ None or very few participated ☐ Some participated ☐ All or nearly all participated	☐ Very rushed ☐ Somewhat rushed ☐ Not rushed at all

Activity 1 **Pregnancy Prevention Memory Game**	*Activity 2* **Birth Control Methods**	*Activity 3* **Keeping it Real: My Relationships**
☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach
NOTES:	NOTES:	NOTES:

Activity 4 **Session 3 Closure**	**ADDITIONAL Session Notes:**
☐ Taught as described ☐ Taught with changes ☐ Did not teach	
NOTES:	

SESSION 4: Relationships and Drugs and Alcohol

Date(s) Taught:		**Time Started:**	
Number of Students Present:		**Time Ended:**	

Were the participants interested?	**Did the students participate?**	**Did the session feel rushed?**
☐ Not at all interested ☐ Some were interested ☐ All or almost all were interested	☐ None or very few participated ☐ Some participated ☐ All or nearly all participated	☐ Very rushed ☐ Somewhat rushed ☐ Not rushed at all

Activity 1 **You Make the Call: Healthy or Unhealthy?**	*Activity 2* **My Reminder**	*Activity 3* **Who Am I? Drugs Version**
☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach
NOTES:	NOTES:	NOTES:

Activity 4 **Heads Up!**	*Activity 5* **Session 4 Closure**	**ADDITIONAL Session Notes:**
☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach	
NOTES:	NOTES:	

SESSION 5: STIs and HIV

Date(s) Taught:		**Time Started:**	
Number of Students Present:		**Time Ended:**	

Were the participants interested?	**Did the students participate?**	**Did the session feel rushed?**
☐ Not at all interested ☐ Some were interested ☐ All or almost all were interested	☐ None or very few participated ☐ Some participated ☐ All or nearly all participated	☐ Very rushed ☐ Somewhat rushed ☐ Not rushed at all

Activity 1 **Project Prep**	*Activity 2* **You Asked…** **Answers to Your Questions about Drugs and Alcohol**	*Activity 3* **STIs and HIV: Legit or Lie?**
☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach
NOTES:	NOTES:	NOTES:

Activity 4 **Where Would I Go?**	*Activity 5* **Session 5 Closure**	**ADDITIONAL Session Notes:**
☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach	
NOTES:	NOTES:	

SESSION 6: Role-Play and Project Prep

Date(s) Taught:		**Time Started:**	
Number of Students Present:		**Time Ended:**	

Were the participants interested?	**Did the students participate?**	**Did the session feel rushed?**
☐ Not at all interested ☐ Some were interested ☐ All or almost all were interested	☐ None or very few participated ☐ Some participated ☐ All or nearly all participated	☐ Very rushed ☐ Somewhat rushed ☐ Not rushed at all

Activity 1 **Review of YOU**	*Activity 2* **Decisions, Decisions!**	*Activity 3* **Practice in Pairs**
☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach
NOTES:	NOTES:	NOTES:

Activity 4 **Role-Playing *Take 2*** ***Think about YOU*** ***Make a Move → Review***	*Activity 5* **Project Prep**	*Activity 6* **Session 6 Closure**
☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach
NOTES:	NOTES:	NOTES:

ADDITIONAL Session Notes:

SESSION 7: *O2T* Projects and Closure

Date(s) Taught:		**Time Started:**	
Number of Students Present:		**Time Ended:**	

Were the participants interested?	**Did the students participate?**	**Did the session feel rushed?**
☐ Not at all interested ☐ Some were interested ☐ All or almost all were interested	☐ None or very few participated ☐ Some participated ☐ All or nearly all participated	☐ Very rushed ☐ Somewhat rushed ☐ Not rushed at all

Activity 1 **Presentation Time**	*Activity 2* **Post-Tests**	*Activity 3* ***O2T* Closure, 3-2-1**
☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach
NOTES:	NOTES:	NOTES:

Activity 4 ***O2T* Closure:** **Your Choice to Share**	*Activity 5* **Presentation of Completion Certificates**	**ADDITIONAL Session Notes:**
☐ Taught as described ☐ Taught with changes ☐ Did not teach	☐ Taught as described ☐ Taught with changes ☐ Did not teach	
NOTES:	NOTES:	

Appendix 2: Pre/Post-Test

Pre / Post (**Circle one**)	Educator:	School/Organization:
Date: _____	Are you: Male O Female O	Age:_____
Race/ethnicity (Select all that apply): O White/Caucasian O Black/African American O Asian O American Indian/Alaska Native O Latino/Hispanic O Native Hawaiian/Pacific Islander		

Please fill in the circle that best matches how you feel about the statement below.

		Yes	No	Don't Know
1.	Can a girl get pregnant the first time she has sex?	O	O	O
2.	Can you get STIs from having oral sex?	O	O	O
3.	Can you get HIV from kissing someone?	O	O	O
4.	Can a person have an STI and not know it?	O	O	O
5.	Can you say "no" to sex if you love that person?	O	O	O
6.	Can some drugs kill you, even if you only use them one time?	O	O	O
7.	Can wearing a condom keep you from getting an STI or HIV?	O	O	O

		Strongly Agree	Agree	Disagree	Strongly Disagree	Don't Know
8.	In the past month, I have had sex (vaginal, anal, or oral.)	O	O	O	O	**O**
9.	Getting pregnant or getting someone pregnant at my age would make it more difficult to achieve my goals.	O	O	O	O	**O**
10.	Having sex would be a good way to keep a boyfriend or girlfriend.	O	O	O	O	**O**
11.	I am too young to do drugs or drink alcohol.	O	O	O	O	**O**
12.	I know how to use a condom properly.	O	O	O	O	**O**
13.	I feel that I have a number of good qualities.	O	O	O	O	**O**

Please go to the next page!

		Strongly Agree	Agree	Disagree	Strongly Disagree	Don't Know
14.	I would insist that my partner or I use a condom every time we had sex.	O	O	O	O	**O**
15.	If a person has unprotected sex one time, s/he can get an STI or HIV.	O	O	O	O	**O**
16.	During the past month, I drank alcohol, used marijuana, cocaine, or other street drugs.	O	O	O	O	**O**
17.	I feel able to say "no" to a friend who wants me to do something that I shouldn't do.	O	O	O	O	**O**
18.	Not having sex is the best way to avoid getting HIV or an STI.	O	O	O	O	**O**
19.	I know who/where to go to in my community when I need help or need to talk.	O	O	O	O	**O**
20.	I plan to wait until I am older to have sex.	O	O	O	O	**O**

Thank you for taking the time to complete this survey!

Appendix 3: Facilitator Instructions for Test Administration

Instructions for Student Pre-test and Post-test Administration

1. Distribute one test form to each student.

2. Instruct students to circle "Pre" (if you are administering the pre-test) or "Post" (if you are administering the post-test) on the upper left hand corner of the test form.

3. To begin administering the test, read the following statement:

   "Each of you now has a test form in front of you. I am going to read each question on the test aloud to you and give you a few seconds to mark down your answers to each question. Your answers on this test will be completely anonymous, which means that you will be the only person who will know what you wrote down. No one else will know which test you filled out, so please be completely honest with your answers. If you don't know what a word means or don't understand a question, just raise your hand and I will help you. When we are finished, I will walk around with an envelope, and you will put your completed forms in the envelope. This way I won't be able to see your answers because you will be putting the forms in the envelope yourselves. If you have any other questions, raise your hand and I will help you. Let's get started."

4. Read the first item aloud and pause for the students to enter their responses.

5. Continue until you have read all of the items. Wait until each student has finished marking his/her test form to begin collecting the completed forms. Assist any students who have questions individually.

6. Walk through the class with the open envelope and have the students place their completed forms in the envelope themselves. DO NOT collect the forms yourself in order to maintain confidentiality.

7. Seal the completed test forms in the envelope and submit to FHCCP's *O2T* Evaluator.

Appendix 4: Parental Opt-Out Permission Form

Dear Parent/Guardian(s):
The Family Health Council of Central Pennsylvania (FHCCP) is offering the opportunity for your child to participate in a newly revised version of *On Second Thought (O2T)* and/or *the Teen Game Plan or Your Game Plan programs*. *O2T* is a curriculum that has been implemented by FHCCP in Central Pennsylvania for over five years. *O2T* highlights the relationship between six topics as they relate to making healthy and informed decisions regarding sexual behaviors.
The six *O2T* topics include:

1. **Self-Esteem**
2. **Puberty**
3. **Teen Pregnancy Prevention**
4. **Healthy Relationships**
5. **Drugs and Alcohol**
6. **STDs, HIV, and AIDS**

Individual Project/Presentation

Your child has been selected to participate in *O2T/TGP/YGP* as part of **Harrisburg School District classes.**
The *O2T/TGP/YGP* Sessions will take place at _____ from _____**-**_____**.**
There are seven session categories, each category session lasts approximately 90 minutes. The final session is an individual project presentation. Details will be given to the students during session one (1) and class time will be allotted to begin work on their project.

On Second Thought emphasizes the importance of resources and support systems in the community. Your child will benefit from your involvement and communication about what s/he is learning in *O2T*. Your child may ask to share some highlights from *O2T* with you or may ask you for assistance with an assignment.
Teen Game Plan/Your Game Plan, is an excellent goal-setting tool for youth to begin communication with you about these important adolescent topics. Many of the topics in *Teen/Your Game Plan* will also be discussed in *O2T*, such as creating goals, recognition of unhealthy behaviors, distinguishing between healthy and unhealthy relationships, deciding when to have sex (or not have sex), contraception; and how these all affect your child's future. We encourage you to read this booklet (provided to student participants) and discuss these topics with your child whether or not s/he decides to participate in *O2T/Game Plan programs*.

Your child's participation is voluntary. S/he will be asked to complete a confidential survey before and after *O2T* to evaluate the effectiveness of the *O2T* program. If you do **NOT** want your child to participate in *O2T/Game Plan*, please sign this letter and return it with your child to **John Harris High School-*Dawn R. Reeves, Coordinator*, Division of Student, Family, and Community Support** by **October 19, 2012.**

If at any time you have questions about *O2T, TGP/ YGP* or would like to see a copy of the curriculum, please contact FHCCP's Outreach & Education Coordinator, **Roxanne J. Carrell,** at **717-243-0523 x 1308 or toll free 800-882-8812.**
Thank you in advance for your support.
Sincerely,

Roxanne J Carrell

Roxanne J Carrell, Community Outreach/Education Coordinator

On Second Thought (O2T) ___No, my child may not participate in *On Second Thought (O2T*.)
Teen Game Plan (TGP) ___No, my child may not participate in *Teen Game Plan (TGP)*
YOUR Game Plan (YGP) ___No, my child may not participate in *Your Game Plan (YGP)*

Student's printed name: _____

Parent's signature_____ Date _____