Re-Examining the Evidence for Comprehensive Sex Education in Schools

Part One: Research Findings in the United States

Stan E. Weed, Ph.D.

Irene H. Ericksen, M.S.

2018

The Institute for Research & Evaluation

THE INSTITUTE FOR
RESEARCH & EVALUATION

6068 S Jordan Canal Rd
Salt Lake City, UT 84129
801-414-3443

Re-Examining the Evidence for Comprehensive Sex Education in Schools

Part One: Research Findings in the United States

The Institute for Research & Evaluation
Stan E. Weed, Ph.D. & Irene H. Ericksen, M.S.

Originally published: September 12, 2017
Updated and Revised: May 1, 2018

The Institute for Research and Evaluation (IRE) is a nonprofit research organization noted for its work evaluating sex education programs over the past 25 years. *IRE* has conducted program evaluations for federal Title V, CBAE, and Title XX projects in 30 states, and has evaluated sex education in three foreign countries, in total collecting data from more than 900,000 teens, and conducting over 100 evaluation studies. *IRE* staff members have published articles in professional journals and presented at professional conferences and workshops. Irene H. Ericksen has served on a national panel of consultants to the CDC-supported *Community Preventive Services Task Force* meta-analysis on sex education effectiveness and as a secondary author for the published study on the same topic (2012). Dr. Stan E. Weed, Founder and Director of *IRE*, has served as a national consultant for federal Title XX and CBAE projects, and was a charter member of the *National Campaign to Prevent Teen and Unplanned Pregnancy*. He has been invited to provide expert testimony about sex education to state legislative bodies, the U.S. Senate, the U.S. House of Representatives, and the White House.

Contents

Abstract

Purpose. *To evaluate the research evidence for U.S. school-based comprehensive sex education (CSE)—instruction on contraception and abstinence within the same sex education program—according to standards derived from the field of prevention research, in order to identify evidence of real effectiveness.*
Methods. *We surveyed the studies contained in three authoritative research reviews of U.S. sex education effectiveness: two sponsored by the U.S. federal government (the Teen Pregnancy Prevention evidence review and a meta-analysis study supported by the Centers for Disease Control and Prevention), and one conducted for the United Nations. These reviews have screened several hundred sex education studies for research quality and reported results for the studies of adequate rigor. We examined the 60 studies of U.S. school-based CSE programs found therein which met that test, and evaluated their outcomes according to meaningful and recommended criteria of effectiveness: sustained effects (detected 12 months after the program), on protective indicators (abstinence, condom use, pregnancy, and STDs), for the main (intended) teen population, based on the preponderance of research evidence. (Note:* consistent *condom use is necessary to provide significant protection from STDs.)*
Results. *For U.S. school-based CSE programs, we found no evidence of effectiveness at producing sustained reductions in teen pregnancy (0 programs) or STDs (0 programs). There were only a few initial findings of increased teen abstinence (three programs in four studies) or condom use (four programs)12 months after the program, but evidence from multiple replication studies did not confirm most of the original positive results. In fact, two of these studies showed harmful program effects. We found no evidence of effectiveness for CSE's purported dual benefit—there were no sustained increases in both teen abstinence and condom use (by sexually active teens) within the same target population. CSE failure rates at producing sustained effects on targeted outcomes included 88% failure to delay teen sexual initiation and 94% failure to reduce unprotected sex. And five school-based CSE programs produced significant* negative *effects: three increased rates of teen sex, one increased teen pregnancy, and one reduced contraceptive use. In contrast, there were seven school-based abstinence education (AE) programs—the often-mentioned alternative to CSE—that produced sustained (12-month) delays in teen sexual initiation. Also, nine studies tested AE impact on condom use and none found a negative effect, strong evidence that AE does not reduce teen condom use.*
Conclusions. *When considering programs in U.S. school settings, measured by meaningful standards of effectiveness, the claims that CSE has been proven effective and AE is ineffective were not supported by this combined database containing some of the strongest and most recent outcome studies of U.S. sex education, as identified by three authoritative sources.* ***In fact, the research evidence indicates that CSE has essentially been ineffective in U.S. school classrooms and has produced a concerning number of negative outcomes. The evidence for AE looks more promising, enough to justify prioritizing additional research.***

I. INTRODUCTION

The short- and long-term consequences of teenage sexual activity continue to be a blight on adolescent populations worldwide. In the United States, they are occurring at alarming levels, in spite of more than 30 years of prevention efforts. For example, the Centers for Disease Control and Prevention (CDC) refers to Sexually Transmitted Diseases (STDs) as a "hidden epidemic," reporting that "1 in 4 sexually active adolescent females has an STD," and that the STD rates for adolescents in the U.S. are rising.[1] Worldwide, the AIDS epidemic continues, with "young people aged 15–24 account[ing] for 45% of all new HIV infections."[2] Although the U.S. teen pregnancy and birth rates were at an all-time low in 2009, they remain the highest among all developed countries.[3] In addition, early onset of sexual activity has been associated with a decrease in mental/emotional health and an increased likelihood of experiencing sexual violence for adolescents, especially among females and younger teens.[4] Given these continuing harms, a high priority for many youth advocates and public policymakers continues to be 1) to reduce teen pregnancies, 2) to reduce STD and HIV infections contracted by youth, and 3) to influence teens to postpone sexual activity.

Efforts to achieve these goals typically focus on 1) promoting abstinence: the delay of sexual initiation (i.e., the onset of sexual activity) for sexually inexperienced (virgin) teens and the return to abstinence by sexually experienced (non-virgin) teens, and/or 2) promoting condom use and other forms of birth control (e.g., birth control

pills, Long-Acting Reversible Contraceptives or LARCs[5]) by those teens who choose to be sexually active. Sexuality education programs that encourage these behaviors in youth populations are viewed by many as a key preventive mechanism through which the negative consequences of teenage sexual activity can be minimized or avoided.

Sex education programs vary widely in their content, methods, and effectiveness, so the fundamental question becomes: Is there a type of program that is more effective than others at achieving these desired results? Some advocacy groups, health professionals, and government officials have endorsed a strategy that is commonly called "comprehensive sex/sexuality education," or CSE (sometimes called "comprehensive sexual and reproductive health education"). It is referred to as a "sexual risk reduction" (SRR) approach to teen sexual health, as contrasted with the "sexual risk avoidance" approach (SRA) that is foundational to the abstinence education (AE) strategy.

The CSE strategy is typically based on the assumption that a sizable proportion of the teenage population cannot be dissuaded from sexual activity. So CSE proponents advocate that the best protection for these youth will be to teach and promote the use of condoms—which can reduce but not eliminate the risk of both pregnancy and STDs—while at the same time promoting continued abstinence for virgin teens and a return to abstinence for those non-virgins who are willing to do so. Thus, the "comprehensive" rationale for CSE is that it is supposed to protect the full spectrum of teens. In other words, there is a hypothesized "dual benefit" provided by CSE programs: they simultaneously increase risk avoidance (by promoting teen abstinence or a return to abstinence) *and* risk reduction (by promoting teen condom use) within the same population of youth. This constitutes the central rationale for CSE and its purported advantage over other strategies, such as the abstinence-only approach to sex education.

It should be noted that while promoting sexual abstinence is a nominal goal for CSE, the amount of attention it receives in specific CSE curricula varies widely; it is often given little emphasis, or may be defined very narrowly, as meaning abstinence from vaginal intercourse but allowing other forms of genital contact. In fact, some organizations that develop and support CSE programs are known to teach that sexual activity is healthy and positive for adolescents as long as they are "ready" for it, and it is "consensual" and "protected."[6] This would appear to be incongruent with the purported inclusion of an abstinence message as a feature of CSE programs.

The *International Technical Guidance on Sexuality Education*, produced by the United Nations Educational, Scientific and Cultural Organization (UNESCO), recommends that policymakers employ "clear, well informed, and scientifically-grounded sexuality education" that is "based on a rigorous review of evidence on sexuality education programmes." The *Technical Guidance* report further states that programs that "emphasized both abstinence and use of condoms and contraception [have been] effective in changing behavior when implemented in school, clinic and community settings" and that such "comprehensive sexuality education" should "become part of the formal school curriculum." Finally, the UNESCO document emphasizes that this school-centered strategy should give priority to Human Immunodeficiency Virus (HIV) and STD protection, in addition to teen pregnancy prevention.[7] These four elements in the UNESCO recommendations—CSE content, evidence-based, school-centered, and HIV/STD-focused—provide the basis for a thorough examination of the available sex education outcome research with the purpose of addressing the question: How effective are CSE programs in schools? That is the subject of the following analysis.

The present report will cover school-based CSE programs implemented and evaluated within the United States. A second report, forthcoming, will review evidence for CSE school-based programs outside the United States.

II. SUFFICIENT EVIDENCE OF PROGRAM EFFECTIVENESS

Referring to sexuality education programs as "scientifically grounded"[8] suggests that such interventions have produced scientifically valid evidence of real success or effectiveness at lowering teen pregnancy, HIV, or STD rates, or at least at increasing the protective behaviors—abstinence and condom use—that prevent or reduce these problems. This raises the critical question of how program success or "effectiveness" is defined and measured.

Most reviews of sex education outcome research set a high standard for the quality of the research *methods* used by the studies included in their database. This is important, since well-designed and well-implemented studies are necessary to produce findings that are an accurate representation of reality. But many of these same reviews have been less careful or clear about their standards for the *outcome measures* they have used to define program success. This is problematic since it is these outcomes—the effects on adolescent behaviors and health—that are the real world impacts by which an intervention's usefulness must be judged. For example, authors of some research reviews may refer to "evidence of effectiveness"[9] without being clear that they consider a reduction in teenage risk behavior that was detected immediately after the program, but disappeared 10 months after the program, to constitute evidence of program success. And some reviews are not adequately transparent about the totality of the evidence of program effectiveness versus ineffectiveness, basing a designation of program effectiveness on one significant effect from a single study while ignoring strong evidence from other studies showing that the same program had no impact or even some negative effects.

Assuming that standards of rigorous study methodology have already been met (so that confidence in findings is high), the broader field of prevention research recommends measuring program effectiveness using certain standards for critical *outcomes*.[10] These standards include a requirement of sustained long-term effects as well as a concern about main effects (on the main or intended population) versus subgroup effects. For example, "sustained impact," defined as "at least one year beyond treatment" is required by the *Blueprints Programs* in order for an intervention to be designated as an effective or model program.[11] *The Society for Prevention Research* (*SPR*) articulates the distinction between program "efficacy" and "effectiveness." The latter requires higher standards than the former. *SPR* defines *efficacy* as the ability of a program to provide some "beneficial effects … under optimal conditions of delivery," and *effectiveness* as the repeated demonstration of positive effects under "real-world conditions."[12] However, to meet even the lower standard of *efficacy*, *SPR* requires evidence from at least two good studies, "a consistent pattern of non-chance findings in the desired direction … there must be no serious negative (iatrogenic) effects on important outcomes," and at least one study showing long-term outcomes measured "at an appropriate interval beyond the end of the intervention."[13] According to *SPR*, *effective* programs must meet these standards for *efficacy* as well as show repeated replication of long-term effects in real-world conditions. Moreover, they consider evidence of effectiveness a prerequisite for a prevention program's dissemination.[14]

Influenced by these entities and their standards, and more than 25 years of experience evaluating school-based sex education programs, the *Institute for Research & Evaluation* has identified five key criteria for evidence of program effectiveness. The first three pertain to the *strength* of a program's outcomes. The latter two have to do with the *quantity and objectivity of the research evidence* about those outcomes. Meeting these meaningful and recommended standards would establish sufficient empirical grounds for evidence of program effectiveness.

A. Impact on Protective Indicators. Given the worldwide epidemic of STDs among young people, sex education programs should not be deemed "effective" unless they increase protection from HIV and STDs, not just from pregnancy alone. That is, they should produce increased rates of either *sexual abstinence* or *consistent condom use* (i.e., using a condom with every act of sexual intercourse). *Consistent* condom use is necessary because STD transmission can occur in one sexual contact and some studies have found that *non-consistent* condom use provided inadequate STD protection or resulted in higher rates of STDs.[15] (Note: Even consistent condom use does not provide the 100% protection from STDs afforded by abstinence,[16] nor prevent the increased emotional harm and sexual violence associated with teen sex.[17] And measuring rates of condom use at last intercourse does not constitute an adequate measure of consistent condom use.) In the body of evidence reviewed here, quite a few studies do not even measure condom use, or they only measure "contraception" which can mean *either* condom use *or* other birth control methods—birth control pills, LARCs, etc. Unfortunately, these latter pregnancy prevention methods provide no protection from STDs or HIV and some may even cause harm.[18] Another commonly used program outcome is to ask teens if they have had "unprotected sex," where a "no" response means they have either used any one of these *contraceptive* methods *or* have been abstinent, without specifying whether the protective behavior employed was abstinence, use of condoms, or use of other types of contraception. Combining these three very different behaviors into one measure by asking students if they have had unprotected sex can make it difficult to determine what the program's protective effect really is, whether it

protects teens from STDs and HIV through increased abstinence or condom use. For this reason, neither the outcome "increased contraception," nor the outcome "[reduction in] unprotected sex" are considered by this review to be adequate measures of program effectiveness. In spite of this inadequacy, we will report on "unprotected sex" when it appears to be a type of surrogate measure for condom use.

B. Sustained (12-month) Post-Program Results. In keeping with standards from the field of prevention research, a program's behavioral impact should last for a sustained period after the end of the intervention. Consistent with several reputable prevention agencies, we define a sustained or long-term effect as at least 12 months following program participation.[19] This is especially meaningful for school-based programs, where another "dose" of the program may not be delivered until a year later, during the following school year, if at all. Thus, a school-based program that produces positive behavior change three or six months afterward, but not when measured at the 12-month follow-up should not be considered effective, and a research study that does not measure this sustained long-term effect has not produced sufficient evidence of a school-based program's effectiveness.

C. Main Effects for the Target Population. The program should produce "main effects"—positive results for the intended/targeted population as a whole and not just for a segment or subgroup of that population (e.g., should affect both boys *and* girls, if both are participating in the program). Most importantly, a program that has produced significant *negative* effects for a substantial subgroup of the intended population (such as boys, or those already sexually active), should not remain on a list of "effective" programs.[20]

D. Based on the Totality of Evidence. The designation of a prevention program as "effective" should take into account the preponderance of evidence about that program's impact. The program should produce "consistent positive effects … [and] no serious negative (iatrogenic) effects on important outcomes"[21] both within the same study and across multiple evaluation studies. Some evidence reviews will report a program to be effective if they can find one positive effect within a single study while ignoring null effects on other more important outcomes in the same study, or evidence from independent replication studies that have found no effect or even negative effects. For example, the U.S. government's signature list of pregnancy prevention programs has included an intervention that produced positive effects in one impact study but no effects in three other rigorous replication studies, and a negative effect in a fourth study.[22] Yet the field of prevention research recommends that positive evidence from multiple studies, without negative effects, be produced before a program is considered effective.[23]

E. Data from Independent Evaluators. The *Society for Prevention Research* (*SPR*) reports that, on average, the findings of prevention program studies are more positive if the study is conducted by the program developer than by an independent evaluator not affiliated with the program.[24] This suggests that an automatic bias or conflict of interest may often occur. Even with rigorous study quality, research results may not be free of this built-in bias that can affect the study findings in subtle ways. When such a study constitutes the sole source of evidence of effectiveness, it calls into question the designation of "effective."[25] *SPR* recommends that program effectiveness should not be founded on evidence produced largely by program developers.[26]

It is not difficult to find sex education programs that have only produced results on less-protective outcomes, or for short durations, only for subgroups of the intended population, or based on a single study conducted by the program's developer and/or implementer. While such outcomes can identify programs that may have potential, they do not constitute sufficient evidence of effectiveness to justify widespread dissemination in school classrooms nor financial support using public tax dollars. In fact, when the totality of evidence for a specific program is examined in detail, such positive outcomes may be offset by countervailing evidence of null or negative effects that would warrant its elimination from lists of "effective" programs.

III. EFFECTIVENESS FOR SCHOOL-BASED COMPREHENSIVE SEX EDUCATION

A high school, middle school/junior high, or elementary school is the setting in which many CSE interventions occur. It is a venue where sex education programs can reach large numbers of their target audience in relatively convenient and cost-effective ways. Perhaps for these reasons, schools tend to be the venue of choice (as in the UNESCO recommendation quoted above) and the focus of the public policy debate about prevention. Our review of sex education effectiveness was conducted with the aim of informing public policy, and for this reason we focused solely on studies of CSE programs that are implemented in school settings.

We define "school-based" sex education as programs that serve a typical school population or recruit participants from such, are held at a school in a classroom-type setting (including after school or on Saturdays), use a curriculum delivered by teachers or facilitators, and can be implemented at most schools. By contrast, clinic or community-based programs often serve unique populations and use methods not easily replicable in schools. Not included in our school-based category are service-learning programs that occur *primarily* in community agencies and settings, and multi-component after-school youth development programs with community and/or summertime components that cannot be implemented mostly within a school classroom setting and methodology. (A prominent example of youth develop programs is the *Children's Aid Society (CAS) Carrera* program.) Within the school-based category, however, we distinguish between two very different types of programs: school-*presented* versus school-*recruited* programs. The first are interventions that can be presented in school classrooms and/or assemblies during the regular school day, and are aimed at a school-wide population (i.e., not a recruited or self-selected subgroup). The second type of intervention recruits participants from within the school (thus they may be different from the general student population). The program is conducted for these recruits after school or on Saturdays (usually in small groups of six to eight participants), and the recruits are often paid to participate. Our review examines studies of both types of school-based programs and distinguishes between these two types of programs in the reporting of findings.

Using the five criteria described above to assess a prevention program's *results*, it is possible to determine which and how many school-based sex education programs have met *these meaningful and recommended* criteria for effectiveness. This report applies these standards to the most recent and best outcome studies of CSE programs in the U.S. The *studies* canvassed in this review have been accepted by other reviewers—that is, not selected by this paper's authors—as meeting sufficient standards for *research quality* (the external reviewers will be mentioned below). Thus, the accuracy of their findings is considered as a given unless otherwise indicated. But a key feature of our review is that the facts and conclusions reported here are derived from our close examination of the original research studies themselves, not by reading the summaries or conclusions of other reviewers. In addition, when a particular sex education program has been evaluated by more than one research study, the findings of all good studies pertaining to that program have been used to inform our conclusions.

IV. THE DATABASE: THREE SCIENTIFIC RESEARCH REVIEWS

The source of empirical evidence about U.S. sex education effectiveness recognized by most policymakers is the universe of outcome studies that have been conducted since about 1990 on sex education prevention programs in the United States, a pool of several hundred studies. This database has been reviewed and sifted by many reputable scientific entities, which have then summarized the results of the studies that met their standards for acceptable research quality. Among such entities are three authoritative governmental agencies: the U.S. *Teen Pregnancy Prevention* program (*TPP*), the U.S. *Centers for Disease Control & Prevention (CDC)*, and the *United Nations Educational, Scientific, and Cultural Organization* (*UNESCO*). Each of these agencies has conducted an extensive review of all the credible studies of CSE conducted in the U.S. during that time frame. Moreover, each has claimed that CSE has shown evidence of effectiveness sufficient to recommend it as a prevention strategy. Because of the prominence of these three entities, and because the included studies met standards for adequate research quality, we chose these reviews as the database for our analysis, in order to evaluate some of the best evidence upon which statements of CSE effectiveness have been based. Since our focus was school-based programs (for reasons stated

previously), we examined the studies of school-based CSE found in these three reviews, applying the meaningful and recommended standards of effectiveness described above to each study's results.

A. U.S. Department of HHS Teen Pregnancy Prevention Evidence Review. As part of the U.S. government's *Teen Pregnancy Prevention* (*TPP*) program, authorized in 2009 by the Office of Adolescent Health (OAH) within the U.S. Department of Health and Human Services (DHHS), a review of the existing sex education research was conducted. That review, overseen by *Mathematic Policy Research*, constitutes one of the most rigorous and current aggregations of research evidence on sex education outcomes extant today. The initial *TPP* Evidence Review examined the sex education research from the prior 25 years, canvassing approximately 600 studies, using standards of research quality to identify the best evidence for program effectiveness available to date. Out of these hundreds of studies, their original review identified only 28 prevention programs described as showing "evidence of effectiveness in reducing teen pregnancy, sexually transmitted infections, and associated sexual risk behaviors" as defined by *TPP* reviewers.[27] (Not represented on that list were the hundreds of studies, out of the original 600, which met the standard for research *quality* but did not have any positive *outcomes*. Thus the picture presented by the *TPP* list is skewed in this sense. It does not reveal the very low success-to-failure ratio overall for the many sex education programs reviewed.) The initial review was updated in 2015–2018, when the originally selected body of studies was supplemented with a subsequent round of outcome studies, including replication studies of some programs identified in the initial round and several new programs being tested for evidence of success.[28] This combined *TPP* database contains 36 studies of 18 *school-based* CSE programs. Included are the original evaluation studies for each school-based CSE program on the original *TPP* list,[29] as well as studies of the school-based CSE programs evaluated in the 2015–2018 *TPP* evidence review (some of which did not qualify for inclusion on the *TPP* list). We have examined this evidence study by study, evaluating the data according to the standards of effectiveness in Section II above.

B. **CDC-Supported Meta-Analysis of Group-Based Teen Pregnancy, HIV, & STD Prevention Programs in the U.S.** The *Community Preventive Services Task Force* operates under the auspices of the U.S. Department of Health and Human Services (DHHS) through support from the *Centers for Disease Control and Prevention (CDC)*. In 2008, the *Task Force* initiated a study of "The Effectiveness of Group-Based Comprehensive Risk Reduction and Abstinence Education Interventions to Prevent or Reduce the Risk of Adolescent Pregnancy, HIV, and STIs." The database included outcome studies from the prior 20 years that met the *Task Force's* standards for research quality, and included 24 studies of school-based CSE interventions. (These studies were selected without regard to the finding of positive program impact, so they give a more realistic picture of the general success-to-failure ratio.) The study concluded that the CSE strategy was generally effective "across a range of populations and settings … [including] both … school and community settings."[30]

C. **UNESCO International Technical Guidance on Sexuality Education, 2009/2018.** In 2009, the *United Nations Educational, Scientific, and Cultural Organization* (*UNESCO*) published an international review of the impact of sexuality education programs on the sexual risk behavior of young people. It surveyed outcome studies in the United States, "other developed countries," and "developing countries," screened them for research quality, and summarized the results. An updated review was published in 2018 which included "22 rigorous systematic reviews and 77 randomized controlled trials … in a broad range of countries and contexts." The 2009 review claimed that "the evidence for the positive impacts [of CSE programs] on behaviour is quite strong," that CSE works equally in school and community settings, has been consistently validated by replication studies, has shown

effectiveness at increasing both teen abstinence and condom use within the same program, and has not increased adolescent risk behavior. The updated review reaffirms the 2009 conclusions, asserts that the evidence base for the effectiveness of school-based CSE "continues to grow and strengthen," and concludes, "programmes that promote abstinence-only have been found to be ineffective" while "programmes that combine a focus on delaying sexual activity with content about condom or contraceptive use [i.e., CSE] are effective."[31] The combined UNESCO database contains 23 studies of U.S. school-based CSE.

As previously stated, our review involved analysis of the individual research studies identified by these three entities as being of sufficient quality for inclusion in their evidence base. It should be noted that there is considerable overlap in the lists of studies included in each of these three reviews. The net result is a set of 60 studies of adequate scientific rigor, evaluating 40 different school-based CSE programs in the U.S. Table 1 summarizes the results for these 60 studies, listed alphabetically by program, and indicates for each study which of the three entities included it in their review (*TPP*, *CDC*, *UNESCO*, or a combination of the three). The color key at the bottom of Table 1 will provide the reader a color map or visual representation of the evidence of CSE effectiveness relative to lack of evidence or evidence of failure.

Our findings are detailed in Section V, below, with Part A presenting evidence of CSE program effectiveness, Part B presenting evidence of CSE program failure, Part C giving a brief comparison of the evidence for school-based CSE and abstinence education (AE) programs, and Part D providing a separate summary of the findings for the U.S. *TPP* list of "evidence-based" CSE programs in U.S. schools.[32]

V. OUTCOMES OF U.S. SCHOOL-BASED CSE

A. Evidence of Effectiveness: U.S. School-Based Comprehensive Sex Education

The 60 studies of 40 school-based CSE programs show very little evidence of CSE success at producing sustained effects (12-months after the program) on important protective outcomes (increased teen abstinence or condom use, or decreased teen pregnancy or STDs) for the targeted adolescent populations.

1. Teen Abstinence

Three of the school-based CSE programs showed some initial evidence (in four studies) of a sustained increase in teen abstinence. However, these findings were not supported in multiple replication studies, including one that found a negative outcome.

a. Sustained (12-month) delay of teen sexual initiation/onset (the most protective behavior)
   - Of the 32 school-based CSE studies that measured this outcome 12 months after the intervention, three programs showed some initial evidence (in four studies) of sustained delays in sexual initiation for the intended population: *It's Your Game: Keep It Real, Postponing Sexual Involvement,* and *Reducing the Risk*.[33]
   - However, research evidence from multiple studies of each program contradicted those initial positive results (see Section A5, *Evidence from Replication Studies of School-Based CSE Programs*), including an independent replication study of *It's Your* Game that showed a significant *increase* in teen sexual initiation for program participants.[34]

b. Shorter-term effects on teen sexual initiation
   Three school-based CSE studies produced delays in teen sexual initiation lasting more than six months but less than 12 months, post-program:

- *Get Real* produced modest but significant delays in sexual initiation (a 15%-16% reduction) six to nine months after a program that spanned sixth to eighth grade, but the study did not measure whether effects lasted until ninth grade.[35]
- A second study of *It's Your Game: Keep It Real* by program developers reduced teen sexual initiation 10 months after the program, but the effect was not detected at the 24-month follow-up, and an independent replication study found a significant *increase* in sexual initiation.[34]
- *Healthy Oakland Teens* delayed sexual initiation for a period that ranged from 8 to 11 months after the program.[36]

c. Other *less-protective* measures of reduced sexual activity
- Seventeen of the 60 school-based CSE studies measured 12-month reductions in "sex in the past 3 months" or frequency of sex (this is a move in the direction of abstinence vis-á-vis reduced sexual activity). However, 12 months after the program, only two studies found a positive result, and one showed a negative outcome.[37]
- Three school-based CSE programs produced a sustained 12-month reduction in number of sex partners,[38] a behavior that still leaves teens exposed to STDs and pregnancy, and requires consistent and correct condom use to reduce risk.[39]

2. Condom Use by Sexually Active Teens

One school-based CSE program reported a sustained improvement in *consistent* condom use, but an independent replication study found no positive effects and significant negative effects. Three other programs produced a 12-month increase in *frequency* of condom use (a less-protective behavior), but studies that replicate these results were not available.

a. Consistent Condom Use (the most-protective condom behavior)
- Of the six studies that measured consistent condom use (CCU) 12 months after the intervention, only one school-based CSE program, *¡Cuídate!*, produced a sustained increase for the target population of teens, in a study by the program's developers.[40]
- Notably, this finding of a 12-month improvement in CCU seemed to be undermined by data from the same study. (See Section A5, *Evidence from Replication Studies of School-Based CSE Programs*.)
- However, an independent replication study of *¡Cuídate!* found no short-term or sustained effects on teen condom use and a significant negative effect—an increase in sexual activity.[41] (See Section A6, *Negative Program Effects by School-Based CSE*.)

b. Other Measures of Condom Use (frequency, use at last intercourse, etc.)
- Of the 12 school-based CSE studies that measured a sustained (12-month) effect on less-protective measures of condom use, such as *frequency*, three programs found significant improvement 12 months after the program in studies by the program's developers (*HIV Prevention Interventions, Safer Choices,* and *Making Proud Choices*).[42]
- However, replication studies to test these initial results are not available.

c. Unprotected Sex

Some CSE studies report on the outcome "unprotected sex," a measure usually obtained by asking teens if they have had sex without a condom or effective means of birth control. This outcome is usually not a clear indicator of teen risk behavior or level of protection. (See Section II, Item #1, "Impact on Protective Indicators.") However, it can serve as a kind of surrogate indicator for program impact on teen condom use.
- Sixteen school-based CSE studies measured a reduction in unprotected sex 12 months after the program and only one, *¡Cuídate!,* found a significant effect. However, a replication study of

the same program by an independent evaluator found it increased teen sexual activity for major subgroups of the program participants.[43]

- Another program (*It's Your Game: Keep It Real*) found a decrease in unprotected sex 10 months after the program, but the effect had dissipated at the 24-month follow-up.[44] The 2012 study also produced evidence suggesting the program *increased* sexual initiation for its male participants after 10 months. (See Section A5, *Evidence from Replication Studies of School-Based CSE Programs*.) And a subsequent study by independent evaluators showed a statistically significant *increase* in sexual initiation for program participants after 12 months.[45]

3. Biological Outcomes: Teen Pregnancy and STDs

Few school-based CSE programs measured teen pregnancy or STDs, and none demonstrated effectiveness at reducing these outcomes.

a. Teen Pregnancy

- Ten of the 60 CSE school-based studies measured the outcome of pregnancy (six measured 12-month effects): none showed reductions 12 months after the program.
- One program (*Teen Outreach Program* or *TOP*) showed effects at the end of a nine-month program, as found in two studies. In one, the effect dissipated 10 months after the program; in the other it was not measured beyond the end of the program. However, another study of *TOP* found a negative effect—an increase in teen pregnancy for the girls in the program.[46]

b. STDs

- Only two of the 60 school-based studies measured program impact (of any duration) on STD infection and neither found any significant effect.

4. The Intended "Dual Benefit" of CSE: Impact on Both Abstinence and Condom Use

The school-based programs in this database did not demonstrate effectiveness at achieving the purported dual benefit of CSE, that is, increasing teen abstinence while simultaneously increasing teen condom use for sexually active teens within the same program.

In theory (according to CSE proponents), there is a dual benefit that constitutes the advantage of CSE programs over AE programs: that they simultaneously increase risk avoidance (by delaying sexual initiation for sexually inexperienced teens and promoting a return to abstinence for the sexually experienced) *and* reduce sexual risk for teens who remain sexually active (by increasing condom use), all *within* the same population of youth.

a. Twenty of the 60 school-based CSE studies measured sustained (12-month) effects on *both* abstinence and condom use (by any measure—whether consistency of use, frequency of use, or use at last sex), and none produced significant effects on both outcomes simultaneously in the same target population.

b. Five school-based CSE programs (in six studies) achieved this "dual" benefit if counting less-protective indicators, or effects on subgroups of the population, and/or for a shorter duration (e.g., three months).[47] However, two of these programs were found in replication studies by independent evaluators to produce significant *negative* effects on program participants (*¡Cuídate!* and *It's Your Game: Keep It Real*—see Section A6, *Negative Program Effects by School-Based CSE*), and two did not measure sustained effects.

5. Evidence from Replication Studies of School-Based CSE Programs

The pattern of evidence from replication studies of school-based CSE programs in this database was not favorable when measured by meaningful criteria and including studies by independent evaluators.

The results for the school-based CSE programs with multiple outcome studies are summarized below.

a. *Reducing the Risk*

Out of eight different studies, there appeared to be more evidence of failure—findings of no effect—than evidence of success for Reducing the Risk *in school settings*.

This database contained eight studies of *Reducing the Risk (RTR)* in school classrooms. All of these measured teen abstinence and condom or contraceptive use as potential program outcomes. (See Table 1 for detailed findings.)

- Out of eight school-based *RTR* studies, only one (a modified version of *RTR*) produced credible evidence of sustained main effects on any protective outcomes: a reduction in teen sexual initiation and in number of sex partners. The study found no positive effects on teen condom use.[48]
- Three other *RTR* studies reported sustained effects on teen abstinence that were based on questionable scientific evidence:
    - An initial study by program developers reported a long-term (18-month) reduction in teen sexual initiation but no effect on contraceptive use.[49] However, the abstinence effect did not hold up in the more rigorous logistic regression analysis and was not recognized by the U.S. *TPP* review as a significant finding.[50]
    - Another *RTR* study found a long-term reduction in teen sexual initiation but no overall effect on contraceptive use. However, this study had serious methodological problems (58% attrition, small sample, no statistical control for existing pretest differences) that call into question the validity of the findings.[51]
    - A third study, actually two studies in one, tested two different versions of *RTR* against each other and a control group.[52] There were no program effects for either of the two versions of *RTR* compared to the controls, but the authors combined the samples of the two different *RTR* programs and reported a significant program effect on sexual initiation compared to the control group. However, this "combined" effect appears to be an artifact since it did not occur in the real world (no adolescent received both versions of *RTR*). Moreover, since the two *RTR* programs were different enough to test against each other (apples and oranges) it does not seem appropriate to combine them and count this as evidence of an *RTR* effect. (The *TPP* website reports a null effect for this outcome in one data table and a positive effect in a different data table.[53])
- Among eight school-based *RTR* studies, there were no sustained 12-month main effects on any other important indicators, including condom/contraceptive use, unprotected sex, or pregnancy.
- Four of the *RTR* studies found no main effects at all, even of short-term duration.[54]
- Finally, there was no evidence for the intended "dual" CSE benefit of increasing both teen abstinence *and* condom use by sexually active teens within the same study population.

b. *It's Your Game: Keep It Real*

There is more evidence of program failure for It's Your Game: Keep It Real (IYG)*—findings of no impact or negative impact—than evidence of program success. In fact, given the evidence for negative impact, IYG appears as likely to harm as to benefit adolescents in school populations*.

- The initial study by the program's developers reported a main effect on teen sexual initiation (defined in this study as the combined onset of anal, oral, and vaginal sex) 12 months after the program, but the effect was not statistically significant for males or for vaginal sex measured separately, and there were no significant program effects on condom or contraceptive use.[55]
- A second set of two studies by the program's developers reported a significant impact on teen sexual initiation and on a combined measure of condom use and abstinence, both at the 10-month follow-up but not the 24-month follow-up. (Effects were found for anal sex but not overall sexual initiation at 24 months.) However, like the first study, the 10-month effect on abstinence was not statistically significant for males, and in this case it was in the negative

direction, *suggesting an increase in sexual initiation for male participants* (AOR= 1.33). This, along with the over-representation of females in the analysis (64%), casts doubt on the finding of a significant overall improvement in teen abstinence.[56]

- Another replication study of *It's Your Game by an independent evaluator*, found *a negative effect* on the main population—a substantial and significant increase in teen sexual initiation 12 months after the program for the full sample of participants and no positive impact on consistent condom use or other contraceptive use.[57]
- And another independent replication study found no significant program effects at all for *IYG* after 12 months.[58]

c. *¡Cuídate!*

There is as much evidence of failure—showing no impact or negative impact—as evidence of success for ¡Cuídate!. *The presence of significant negative effects from an independent replication study would seem to outweigh the positive effect on consistent condom use reported in the study by program developers*.

- The original study of *¡Cuídate!* (by the program developers) found no effect on teen abstinence, but reported a 12-month improvement in rates of consistent condom use (CCU) and reduction in the number of sex partners (a less-protective effect).[59]
- This claim of a 12-month program impact on CCU is called into question by data from the same study, wherein a pretest difference, not controlled for, appeared to account for nearly all of the 12-month difference between groups that was attributed to a program effect.
- A replication study by independent evaluators looked at the impact of *¡Cuídate!* in a school classroom setting and found no positive results and significant negative effects on substantial subgroups of participants. (See Section A6, *Negative Program Effects by School-Based CSE*.).[60]

d. *Teen Outreach Program (TOP)*

When looking at the five evaluation studies of Teen Outreach Program (TOP) *in schools, there is more evidence of program failure—findings of no impact or negative impact—than evidence of sustained positive impact*. TOP *has shown no evidence of long-term post-program benefits and the potential to do harm in adolescent school populations*.

The *TOP* is a school-based youth development and service-learning program with a sexuality education component that includes a CSE approach to pregnancy prevention.

- The initial study of the *TOP* measured teen pregnancy at the end of the nine-month program and found a significant reduction for program participants. However, no follow-up measure was taken to test for the duration of this effect beyond the end of the program.[61]
- A recent replication study in Florida schools found positive *TOP* effects on teen abstinence and pregnancy at the end of the program, but these were not sustained 10 months after the program's end.[62]
- A recent study of the *TOP* in Minnesota schools found no significant effects at three or 15 months after the program on any outcomes—teen sexual initiation, recent sex, or unprotected sex.[63]
- Another recent replication of the *TOP* in Chicago found no effect on consistent condom use (the only outcome measured).[64]
- And a recent large multi-site evaluation of the *TOP* in the Northwestern U.S. found a significant *increase* in the rate of pregnancy for females, and no positive effects.[65]

e. *Postponing Sexual Involvement (PSI)*

Three studies (two by independent evaluators) showed little evidence of success for Postponing Sexual Involvement *in school settings*.

- The initial study found a 12-month delay in teen sexual initiation, but it was rated as "a weak design with many problems" by reputable reviewers.[66]
- A subsequent replication study of *PSI* found no sustained effects for the intended population (only short-term subgroup effects).[67]
- A third *PSI* study found no effects on sexual initiation, recent sex, or number of partners, even short-term.[68]

f. *Be Proud Be Responsible (BPBR)*

Three studies (two by independent evaluators) showed no evidence of sustained program effects for Be Proud Be Responsible *in school settings*.

- An initial study by the program developer found a reduction in unprotected sex and anal sex (but not vaginal sex) six months after the program.[69]
- A replication study measured 12-month outcomes for sexual initiation, consistent condom use, and unprotected sex and found no effects.[70]
- An adaptation of *BPBR* found a reduction in unprotected sex at six months but not 12 months after the program, and no impact on teen pregnancy.[71]

g. *The Children's Aid Society (CAS) Carrera Program*

The evidence from six studies of the CAS Carrera *program is not favorable: no sustained post-program effects were measured, and there appears to be more evidence of program failure—both null effects and negative effects—than program success*.

This multi-component positive youth development program is in a different category than the school classroom type CSE programs that are the subject of this report. However, because the program draws its participants from school populations, emphasizes both abstinence and contraception, is on the *TPP* list of evidence-based programs that *can* be implemented in schools, costs nearly $5,000 per student, and has been the subject of multiple replication studies, the outcome evidence is reviewed here. Given that the *CAS Carrera* program is a departure from the school-based typology in this report, the data are not included in Table 1, nor counted in the aggregations of school-based CSE findings.

Six studies of *CAS Carrera* effectiveness have been conducted, four with a randomized design:

- The first study did find some results for girls but not for boys at the end of the three-year program--reductions in sexual initiation and pregnancy. But it found no effect on condom use, and girls in the program were more than twice as likely as those in the control group to use Depo-Provera—a hormonal contraceptive injection—at last intercourse. No measures were taken to determine if these immediate post-program sub-group effects lasted beyond the end of the program.[72]
- A 2009 review by Douglas Kirby of the cumulative outcome evidence from the first four *CAS Carrera* studies found that "One pattern is clear, consistent, and discouraging—none of the four studies found any positive effects on sexual behaviour in young men … In girls, three of the four studies failed to find a significant benefit on current sexual activity or use of contraception, and two [studies] reported significant *increases* in pregnancy rates."[73]
- Two recent replication studies *of CAS Carrera* (2015 and 2016) found no significant effects at the end of the three-year program on rates of teen sexual initiation or unprotected sex (the effect on pregnancy was not measured).[74]

6. Negative Program Effects by School-Based CSE

Five out of the 40 school-based CSE programs represented in this database produced significant negative effects for the main teen population or substantial subgroups—increases in sexual initiation, recent sex, oral sex, or pregnancy. The field of prevention research stipulates that "serious negative effects on important outcomes" should disqualify a prevention program from being designated as "effective." Three of these programs are currently on the U.S. government's *TPP* list of evidence-based programs.[75]

a. ¡Cuídate!
A rigorous replication study of this program in a school classroom setting *by an independent evaluator* (not the program developers) found no positive results and significant negative effects for substantial subgroups: program participants who were sexually active at baseline were more likely to have had recent sex six months after the program, and White participants were more likely to have had oral sex at the six-month follow-up. The study abstract seemed to downplay these negative impacts on important subgroups by stating, "Exploratory subgroup analyses suggest potentially problematic effects for some groups."[76]

b. It's Your Game: Keep It Real (IYG):
A rigorous replication study of IYG *by an independent evaluator* found a negative effect on the main population—a significant increase in teen sexual initiation 12 months after the program, and no positive impact on consistent condom use or other contraceptive use.[77]

c. *Teen Outreach Program (TOP)*:
A replication study of this program *by an independent evaluator* found no positive effects on rates of sexual activity and an increase in the pregnancy rate for female participants at the end of the nine-month program.[78]

d. Healthy for Life
Program participants were significantly more likely to report having sex recently, 24 months after the program.[79]

e. Project SNAPP
Participants had significantly lower levels of contraceptive use, 17 months after the program.[80]

B. Evidence of CSE Failure in School Settings

There was much more evidence of program failure than success for school-based CSE. Failure rates for sustained effects on the most-protective outcomes ranged from 76% to 100%.

Empirical evidence about a sex education program's effectiveness can fit into one of four conditions: 1) *evidence of program success*—the desired outcome(s) were measured, and the results were statistically significant in the positive direction, 2) *lack of evidence of effectiveness*—evidence does not exist about specific program outcome(s) because they were not measured or were measured, and the results were deemed inconclusive, 3) *no effects*—the outcome(s) were measured and the results were *not* statistically significant (i.e., null), or 4) *negative effects*—the outcome(s) were measured, and the results were statistically significant in the wrong direction, indicating a harmful effect. We call these latter two conditions *evidence of program failure*. A sex education program can be said to *not show evidence of effectiveness* because evidence does not exist(the second condition) or because evidence exists but it is *evidence of failure* (the third or fourth condition). But *evidence of program failure*—due to null or negative effects—is a more serious matter than *lack of evidence* of program impact. Table 2 shows numerical values for both situations: the *lack of evidence* about program effectiveness can be seen in the top row indicating the number of studies not measuring the desired outcomes; *evidence of program failure* can be seen in the bottom row showing the proportion of studies measuring an outcome and finding no positive effect. In the previous sections, we focused on the *evidence of program success*. In this section, we summarize the *evidence of program failure* for CSE

programs in U.S. schools—the measuring and finding of no significant, sustained, positive effects on the most-protective outcomes (i.e., findings of null or negative effects) for the target population. It should be noted that the *TPP's* initial evidence review rejected hundreds of CSE studies, many of which were school based, that met *TPP* standards for research quality but which found no positive program effects at all. Thus, the "failure rates" reported here provide a conservative estimate because *they do not reflect the very low success-to-failure ratio overall for the many school-based CSE programs the TPP reviewed and rejected*.

1. ***CSE's Intended "Dual Benefit:" Sustained effects=100% Failure***

    a. Twenty school-based CSE studies measured 12-month changes in both teen abstinence *and* condom use, and none found significant improvements for both. Thus, a 100% failure rate.
    b. Thirty-six CSE studies measured *any* type of dual program benefit—any abstinence and condom increase of any duration, and six studies found significant effects for five programs, an 83% failure. However, two of these programs (representing three of the six studies) also produced negative effects, in rigorous independent replication studies, by significantly increasing teen sexual activity. This would seem to nullify these programs' claim to producing a dual benefit, since abstinence is one of the desired dual benefits. Thus, the net CSE failure rate at producing any dual benefit was 33/36 studies or 92% for school-based programs.

Looking at these dual benefits—abstinence and condom use—separately, gives the following results:

2. ***Teen Abstinence: 88% CSE Failure***

    a. Among school-based CSE programs, 32 of the 60 studies measured program impact on teen sexual initiation for at least 12 months after the program. Only four of these 32 studies, representing three CSE programs, found a significant effect, for an 88% failure rate.
    b. Stated another way, 12% of school-based CSE studies that measured this outcome demonstrated success.
    c. Seventeen of the 60 school-based CSE studies measured reduced "sex in the past three months," or reduced "frequency of sex," movement in the direction of abstinence, 12 months after the program, with two positive results and one negative outcome, an 88% failure rate.

3. ***Teen Condom Use: Consistent Use=No Success; Increased Frequency=76% Failure***

    a. Only six of the 60 school-based CSE studies measured a 12-month effect on consistent condom use (CCU), and only one reported a significant effect. This appears to be a high failure rate, but too few studies exist to estimate a numerical value.
    b. Seventeen of the 60 studies measured a 12-month effect on any indicator of condom use (including CCU, frequency of use, etc.), and four reported a significant effect. This is an overall 76% failure to achieve a sustained improvement in any measure of teen condom use.

4. ***Unprotected Sex by Teens: 94% Failure***

    - Sixteen of the 60 studies measured a 12-month effect on unprotected sex, and only one showed a significant reduction, a 94% failure rate.

5. ***Teen Pregnancy & STDs: No Success***

    a. Only six of the 60 CSE studies measured a 12-month effect on teen pregnancy, none found a positive effect, and one found a negative short-term effect. Thus, there was a general failure on this outcome, but too few studies exist to estimate a numerical value.
    b. Four of the 60 CSE studies measured program effects on teen pregnancy of shorter duration; two studies (both of the *Teen Outreach Program*) found reduced pregnancy immediately following a nine-

month program, which in one case dissipated at the 10-month follow-up measure, in the other it was not measured beyond program end. In another study, the same program *increased* teen pregnancy.

c. Only two studies measured STD effects of any duration, and neither found a significant impact. Thus, there was no evidence of school-based CSE success on this outcome.
d. There is simply a substantial *lack of evidence* about school-based CSE impact on teen pregnancy or STDs.

C. School-Based CSE and Abstinence Education (AE): Relative Evidence

Although AE studies are relatively few in number, there appears to be somewhat better evidence for promoting abstinence through school-based AE than CSE.

The sex education strategy most often mentioned as a counterpoint or alternative to comprehensive sex education is what proponents refer to as "sexual risk avoidance" or "abstinence education" (hereafter "AE"), also referred to by some as "abstinence-only" programs. In contrast to CSE, the AE approach typically teaches youth to abstain from overtly sexual behavior with another person (including vaginal intercourse, oral and anal sex, mutual masturbation, and heavy petting) until they can form a mutually monogamous relationship in adulthood (preferably marriage), as the only way to eliminate risk (rather than merely reduce it) and avoid all the negative consequences of teen sex. Condom use is sometimes addressed in AE programs, but often in terms of its limitations and failure rates; AE does not promote or demonstrate condom or contraceptive use.

A common observation by reviewers of sex education research is the lack of good quality outcome studies of AE programs relative to CSE programs. This is due in part to the fact that the sheer number of studies that have been conducted and published to date is much larger for CSE than AE programs (federal funding for independent outcome studies of AE was cancelled in 2010, ending an opportunity to substantially expand the AE evidence base). In the present database, consisting of studies accepted for sufficient study quality by three credible external reviews, there are 60 studies of 40 school-based CSE programs and 18 studies of 16 school-based AE programs. In addition to the small number of adequate studies, another issue with the AE evidence base has to do with six ostensibly rigorous studies that have serious methodological limitations such that the research design would tend to underrepresent the impact of the AE programs they evaluated. Because none of the six studies found significant program effects they are often cited as evidence of AE ineffectiveness. Unfortunately, their null findings combine to form a faulty evidence base that has weighed heavily in most reviews of AE effectiveness and undermined the case for AE efficacy.[81] For these reasons, we have not produced a detailed summary of the AE research evidence here, nor attempted to draw conclusions based on that evidence. We agree with other reviewers that while there are significant positive findings, the evidence is not of sufficient quantity or quality to draw firm conclusions from the data.[82] Having said that, we will report several trends from the research findings on AE:

First, there appears to be somewhat better evidence in this database for promoting teen abstinence through school-based AE than CSE. As already stated, three school-based CSE programs (in four studies) showed sustained 12-month main effects on teen abstinence (delayed initiation), but multiple replication studies (12 total) showed null or negative effects that seem to outweigh the initial positive findings for these three programs. Conversely, among the 18 school-based AE studies that were of sufficient quality for inclusion in this database, seven programs (in seven studies[83]) showed sustained main effects on teen abstinence. Five of the seven were by independent evaluators. However, only two replication studies have been conducted, one showed promising but inconclusive results, and the other was not confirmatory.[84] More replication studies should be done to verify the initial positive results of these seven studies.

Second, it is important to note that there was strong evidence in this database that contradicts the claim of critics that AE does harm through reducing the use of protection by sexually active teens. Of the nine rigorous AE studies that measured condom use as an outcome, eight found no significant effects, and one showed a significant 12-month improvement.[85] This is compelling evidence that AE does not do harm by causing sexually active teens to *reduce* teen condom use.

Third, there is not adequate evidence about AE impact on pregnancy or STDs—very few studies measured these outcomes, and those that did had some methodological problems. However, the increases in teen abstinence documented in other AE studies would be expected to produce reductions in these outcomes, though unmeasured.

Finally, one of the AE studies found short-term negative effects that disappeared at the longer-term follow-up and were replaced by several sustained positive outcomes.[86]

D. The U.S. *Teen Pregnancy Prevention* List of "Evidence-Based" Programs

The 18 school-based CSE programs designated by the *Teen Pregnancy Prevention* program as showing "evidence of effectiveness in reducing teen pregnancy, sexually transmitted infections, and associated sexual risk behaviors,"[87] provide very little evidence of sustained effects on these outcomes for the intended teenage population. Overall, there is far more evidence of failure than success for these CSE programs. Despite the fewer number of studies, there appears to be promising evidence for the AE programs on the *TPP* list.

As a service to U.S. federal policymakers, in this section we summarize the scientific evidence of effectiveness for the school-based CSE and AE programs that met the United States *Teen Pregnancy Prevention (TPP)* program's criteria for inclusion on its list of evidence-based interventions. The outcome studies that evaluated these school-based programs constitute a subset of the database for the present research review.

It should be noted that while the *TPP* evidence review placed a high priority on the quality of study *methodology*, it had less rigorous standards for the program *outcomes* it used to define effectiveness. These criteria were: to show at least one statistically significant favorable effect—of any duration or for any subgroup—on sexual risk behavior, pregnancy, or STDs. Thus, a program could make the *TPP's* list of programs with "evidence of effectiveness":

- by virtue of just one positive study by the program's developer (for the original 28 programs, only two of the studies were by independent evaluators), while independent studies found null or negative effects,
- by showing only one significant effect on a less-protective outcome (such as reduced number of sex partners) while showing failure to impact the most-protective outcomes like abstinence or condom use,
- without achieving any "main effect" (i.e., impacting only a subgroup of the intended population), or
- without showing a sustained (12-month) effect on any outcome.

Table 3 summarizes research findings for the school-based CSE (Table 3A) and AE (Table 3B) programs designated by the *TPP* website as showing "evidence of effectiveness" (as defined above).

1. Outcomes for the 36 studies of the 18 school-based CSE programs on the TPP list:

a. Teen Pregnancy: None of the 18 school-based CSE programs showed effectiveness at reducing teen pregnancy. While the *Teen Outreach Program (TOP)* reported a reduction in teen pregnancy, it was not a sustained post-program effect, and a subsequent study in a different location found the program actually *increased* pregnancy rates.

b. STD Prevention: None of the school-based CSE studies demonstrated a reduction in teen STDs, in fact, none measured it.

c. Teen Abstinence: None showed effectiveness at increasing teen abstinence. While two of the 36 school-based CSE studies reported a 12-month increase in teen abstinence for the intended population (*Reducing the Risk* and *It's Your Game: Keep It Real*), 10 other studies of the same programs found no such positive effects and one *negative* effect.

d. Consistent Condom Use: None of the 18 school-based CSE programs showed effectiveness at increasing consistent condom use by teens. (*Consistent* use is necessary to provide meaningful protection from STDs.) Although there was one program that reported a long-term effect (*¡Cuídate!*), a subsequent replication study conducted by independent evaluators—not the program's developer—found that the program actually *increased* other sexual risk behaviors, negating the program's claim to effectiveness.

e. Condom Use Frequency (a less protective factor): Two of the 36 studies reported 12-month increases in *frequency* of condom use for the intended population in studies by program developers, however, these results have not been verified in independent replication studies.[88]
f. CSE's Intended Dual Benefit: None of the school-based CSE programs showed effectiveness at achieving the purported dual benefit of the "comprehensive" strategy—increasing both teen abstinence and condom use within the same adolescent population. No program produced sustained effects on both outcomes and the two programs that produced effects of shorter duration or effects on lesser outcomes also produced *negative* effects on other important teen risk behaviors (*¡Cuídate!* and *It's Your Game: Keep It Real*).
g. Negative Effects: Four of the 18 school-based CSE programs evaluated by these 36 studies produced significant negative effects (i.e., increases in sexual initiation, recent sex, oral sex, or pregnancy) for the target population or a substantial subgroup of teens: *CAS Carrera*,*¡Cuídate!*, *It's Your Game: Keep It Real*, and *Teen Outreach Program*.

(See Section A5, *Evidence from Replication Studies of School-Based CSE Programs* for above study details. It should be noted that while the *CAS Carrera* program doesn't fit the school classroom-based typology of the present report—it is usually considered a youth development program—it is included in this summary of *TPP* programs because of its CSE content and the fact that *TPP* reviewers classified it as either school- or community-based.)

2. *Outcomes for the five studies of the five school-based AE programs on the TPP list:*

a. Teen Abstinence: Four of the five AE studies (three by independent evaluators) produced a 12-month increase in teen abstinence. Studies should be done to replicate these initial positive results.[89]
b. Condom Use: Although improving teen condom use is not a goal of AE, it is important to note that three of the five AE studies measured this outcome, and one found a 12-month increase in frequency of condom use. The other two studies found no significant effects, positive or negative. Thus, of the three studies that measured AE impact on condom use, none found a negative effect.[90] This evidence contradicts the claim that AE reduces teen condom use.
c. Teen Pregnancy & STDs: None of the AE studies measured these outcomes, however, programs that increase teen abstinence increase the protective behavior by which teens avoid both of these problems.

VI. SUMMARY

We have surveyed the studies found in three nationally recognized reviews of sex education outcome research—reviews that screened several hundred sex education studies for research quality, and reported the outcomes of those deemed scientifically sound. We examined the 60 studies of school-based CSE that these reviews determined were of adequate quality, and evaluated their outcomes according to meaningful criteria for program effectiveness derived from the field of prevention research: sustained (12-month) effects, on protective indicators, for the main (intended) school population, and based on the preponderance of research evidence (including studies from independent evaluators).

The results paint a markedly different picture than the one depicted in the *UNESCO* report (CSE programs have been "effective in changing behaviour when implemented in school, clinic, and community settings,"[91]), reported by the *CDC* meta-analysis (CSE is effective in "both … school and community settings") or on the *TPP* website ("programs with evidence of effectiveness"),[92] or claimed by some CSE advocates (for example, see *Advocates for Youth*, "Comprehensive sex education has been proven effective … [to] delay onset of sexual activity … and increase condom use"[93]). **For school-based CSE programs, we found a few initial findings of sustained improvement in teen abstinence or condom use followed by evidence from replication studies that did not**

confirm most of the original positive results. There was virtually no evidence of success at reducing teen pregnancy or STDs.

Of particular concern is the dearth of findings of real success by school-based CSE programs at producing sustained improvement on any measure of condom use. None showed effectiveness at increasing *consistent* condom use and only three showed sustained increases in *frequency* of use, a less-protective outcome, in single studies by program developers. This is striking since it is a central purpose of CSE, is one of its main distinctions from AE, and is important for providing even partial protection from STDs for sexually active teens.

Also concerning is the fact that while more than half (36/60) of these CSE studies employed measures that tested CSE's intended dual benefit—simultaneous increases in rates of teen abstinence and condom use (by the sexually active)—there was a startling scarcity of any positive results on both outcomes within the same population. There was no long-term success and only five programs with short-term or lesser effects, two of which also produced other negative outcomes. Again, this is the signature rationale for CSE—that it will effectively increase risk avoidance by promoting abstinence and at the same time reduce risk for teens who decline to be abstinent—and is the advantage it claims over AE. However, there appears to be strong evidence that this is not occurring for CSE programs in school settings and populations.

Finally, and perhaps of most importance, the oft-repeated assertion that CSE programs have done no harm to adolescents is not born out by these research findings. Independent evaluations of five out of 40 CSE programs in schools (six if you include the *CAS Carrera* program for community or school settings) found they produced significant *negative* effects: three increased rates of teen sex, one increased teen pregnancy, and one reduced contraceptive use. Three of these school-based programs are currently on the U.S. Department of Health and Human Services *Teen Pregnancy Prevention* list of evidence-based programs: *It's Your Game: Keep It Real, ¡Cuídate!*, and *Teen Outreach Program*, as is the *CAS Carrera* program for communities or schools.

VII. DISCUSSION

A. The first question raised by these findings is why they differ so dramatically from the common perception that CSE has been proven effective and AE shown to be ineffective and harmful. We suggest several possibilities:

1. Many research reviews by otherwise credible entities have not assessed CSE program *outcomes* by meaningful criteria for program effectiveness. Instead, they have tended to overplay the evidence and accept much lower benchmarks of success, wherein any statistically significant positive change on any indicator for any subgroup or of any short-term duration is called "evidence of effectiveness" for that program, meanwhile ignoring other studies showing null or negative effects. This contradicts principles of program effectiveness from the field of prevention research. At the same time, AE programs have had a higher bar to meet by virtue of measuring sexual initiation—a one-time, all-or-nothing behavior—as the critical outcome, rather than merely measuring sliding scale reductions in frequency of sex or increases in condom use. Furthermore, most AE studies have measured at least a *12-month* duration of effect, which is longer and more difficult to achieve than what has been required of many CSE programs. Thus, this higher bar likely has made it more difficult for AE studies than CSE studies to show statistically significant positive effects. And when held to the same higher standards, the evidence of CSE *ineffectiveness* becomes more clear. As seen in the above analysis, the "low bar" for CSE effectiveness has not been adequate to produce reductions in teen risk behaviors that are sufficient to reduce pregnancy or STDs for program participants. The more stringent effectiveness standards recommended in this report are more likely to identify and/or generate programs that provide real protection for adolescents.

2. Methodological weaknesses in six key AE studies (five conducted by researchers at Mathematica Policy Research, Inc.) appear likely to have underestimated AE program effects, thereby producing questionable study results.[94] However, because these studies have been included in most research reviews and meta-analyses of AE, their null findings have made a large quantitative contribution to the conclusion reached

repeatedly by such reviews that AE programs are categorically ineffective. Thus, these questionable studies have played a major role in creating the pervasive perception of AE ineffectiveness that is refuted by the findings of the credible and more current database reviewed here.

3. Program setting and population are relevant. We have observed that there is a pattern in the sex education outcome research wherein *school-based* CSE programs overall tend to have less positive results than CSE programs in *clinic or community settings*, and that most research reviews tend to blur this distinction. The fact that our review was limited to school-based CSE programs (in the U.S.) has brought this poorer performance into focus. This is important information since these school-based CSE outcome studies occur with the targeted setting and population (adolescents in schools) and strategy (comprehensive sex education) endorsed by UNESCO and by many public policymakers.

4. Sometimes reviews of sex education effectiveness or advocates for CSE make statements that appear to contradict the actual research evidence. For example:
    - A recent report by *Advocates for Youth* states that "No abstinence-only program has yet been proven through rigorous evaluation to help youth delay sex for a significant period of time …"[95] Yet studies of two "abstinence-only" programs have produced significant and sustained delays in teen sexual initiation—one at the 12-month follow-up, and the other 24 months after the program. The studies of both programs were accepted as "evidence of effectiveness" by the *TPP*'s evidence review.[96]
    - A recent research review by the CDC-supported *Community Preventive Services Task Force* concluded that comprehensive risk reduction programs (meaning CSE) were generally effective "across a range of populations and settings…both…school and community settings."[97] However, the detailed results of this meta-analysis study, shared in public settings but not reported in the published research article (they are published in a companion piece in the same journal[98]), showed significantly poorer results for school-based CSE on key outcomes. And the effects of school-based programs were not statistically significant for increasing teen condom use or use of protection, or for decreasing teen pregnancy or STIs. In addition, the effect on pregnancy was in the negative direction, suggesting these programs in schools may have *increased* teen pregnancy. Moreover, nearly one-half (47%) of the 15 school-based CSE studies produced findings suggesting some negative effects on teen condom use.[99] These data present a very different picture than the one depicted by the published report of the study.
    - One review of sex education in schools reported its findings on the effects of "Comprehensive Interventions" as: "Whilst positive changes in reported behaviour were observed in some studies, findings were not consistent enough to draw firm conclusions (Jones et al., 2009a; Kim & Free, 2008; Kirby, 2005, 2007; Underhill et al., 2008; Yamada et al., 1999). Indeed, some studies found improvements while others reported negative or null effects for the same outcome. Health-related outcomes were rarely reported, and when they were, few positive changes were observed (DiCenso et al., 1999; Jones et al., 2009a; Kirby, 2005, 2007; Underhill et al., 2008). One review presented evidence that, in some instances, comprehensive programmes may increase sexual intercourse (Kirby, 2005)" and "It was often not possible to identify … change that could be attributed to exposure to an intervention … positive changes were inconsistent." In spite of these findings, the study *Abstract* asserts that "comprehensive interventions … were found to be effective."[100]

B. A second question worth asking is why are these school-based CSE programs so ineffective, especially compared to programs in other settings? We offer several factors for consideration:

1. First, interventions in clinics and community settings often have a higher-risk population than school-based programs; such teens may be more motivated to learn about and utilize protective measures. In addition, programs in these venues are often able to use methods—such as individual clinical services (e.g., injections of contraceptive hormones), one-on-one counseling and instruction, and regular follow-up phone calls—that are not as easily implemented in school settings and populations.

2. Many of the programs in schools rely heavily on teaching information and skills; some are developed around social learning theories like "The Theory of Reasoned Action" or "The Theory of Planned Behavior." The assumption is that adolescents will plan ahead and apply their new knowledge and skills in rational ways when they find themselves in highly intense romantic interactions. This is related to another possible influence…

3. A biological reality that is not often mentioned in public discourse about sex education is the teenage brain. The science of brain research has reached a consensus over the past 25 years that important regions and functions of the human brain are not fully developed until *after* early adulthood. These include the executive functions of the frontal lobes (governing impulse control, anticipation of consequences, judgment, planning, goal-setting, and prioritizing) and the hippocampal formation and amygdala (areas that mediate motivation, memory, attention, and emotional/affective behavior).[101] According to experts, this means the adolescent brain is physiologically geared for impulsiveness and "risk-taking behavior,"[102] immature processing of information, and failure to anticipate the future impact of behavior, making it "difficult for them to understand and use contraceptive methods effectively and consistently."[103] In other words, adolescents aren't neurologically well-equipped for "reasoned action" or "planned behavior," especially in highly emotional, impulse-driven situations.

4. Related to this is another seldom-mentioned issue: Condom use error and failure can significantly compromise the protective benefits of condom use, and error/failure rates are surprisingly high, even among experienced and motivated *adult* condom users. For example, among 1,973 adults at an urban STD clinic *who were consistent condom users*, 57% of women and 48% of men reported at least one incident of condom use error or failure over a four-month period with condom breakage being the most frequent problem and condom error associated with higher STD levels for men.[104] And in a sample of 102 college women who put condoms on their male partner(s), 30% to 50% (depending on the type of error) reported they had committed a common condom use error at least once in the past three months and 28% reported condom breakage, slippage, or both occurring during sex over the same time frame.[105] We would expect such problems with condom use to be exacerbated in adolescent populations and this may have contributed to the lack of program effects on pregnancy and STDs for the studies in this database that measured these outcomes.

C. A third concern has to do with the finding of negative effects by a significant number of school-based CSE programs. For the three of these programs that are included on the U.S. *TPP* list (giving the appearance of federal endorsement), these negative effects are not readily apparent to someone looking for an effective program on the *TPP* website, and these three programs are implemented in many schools across the U.S.—an unfortunate instance in which the "buyer beware" adage must be applied. It is beyond the scope of this paper to explore what might be causing these negative program effects (which include increased teen sexual initiation, recent sex, oral sex, pregnancy, and reduced contraceptive use). For those interested in pursuing causality, one place to start may be to examine the content of these programs. For example, one of the *TPP* programs that increased the rate of teen sex, *It's Your Game: Keep It Real*, asks seventh and eighth grade mixed-gender classes to engage in role plays that include the following phrases:

- "She is really hot and I've been thinking that maybe it will be OK to mess around a little more than just kissing."
- "I think we should do more than just kissing and touching."
- "I just feel so close to you. That's why I want to have sex with you."
- "If we use a condom, it will spoil the mood."
- "You just need to do it, and then you'll realize sex is no big deal."[106]

Even though these statements are presented as "pressure lines" to be refuted, their mere discussion may suggest to seventh graders that these behaviors are within the normal and accepted range of issues to be negotiated between 12-year-old boys and girls. And this curriculum seems to normalize intimate "touching," which many parents consider sexual foreplay that is inappropriate for young teens. Another CSE program that produced negative effects, *¡Cuídate!* (also on the *TPP* list), uses the following prompt for a discussion with teens as young as 13 years old:

- "What are some of the things that you should consider to help you decide if you are 'ready' for sex?"
- "Possible Answers: If you know this is the right decision for you—now and in the future; if you can talk to your partner about sex; before you have sex, if you know how to protect yourself and your partner; if you can deal with the consequences of having sex—like getting pregnant, an STD or HIV...You shouldn't have sex until you are ready—and until you decide."[107]

It may be hard for parents to imagine their 13-year-old daughter making a mature decision about whether she is "ready" for sex, especially in the face of pressure from an older boy. Thus, school boards, administrators, and parents, may want to investigate whether any negative effects have been caused by sex education programs they are considering and look into the actual content of the program's curriculum, rather than relying on the endorsement of

a federal agency or the program's developers.

D. A final point worth considering is the question of terminology. The sex education strategy examined by this report is typically referred to as "comprehensive sex education" (CSE), "sexual risk reduction" (SRR), or "comprehensive risk reduction" (CRR). Yet the evidence from this database indicates that in U.S. schools these programs have not been effective at achieving the dual benefit from which the term "comprehensive" was originally derived—increasing both teen abstinence and condom use. Similarly, the evidence indicates that these programs have generally failed to "reduce risk" among adolescents in schools. Thus, the findings call into question use of the current labels to describe this strategy.

VIII. CONCLUSION AND RECOMMENDATIONS

We have analyzed the studies found in research reviews by three reputable scientific agencies, containing some of the strongest and most recent evaluations of U.S. sex education programs available—60 studies of school-based CSE that have been screened for research quality by credible entities. Our findings demonstrate that applying a meaningful definition of *effectiveness* to program outcomes and examining the evidence for school-based settings separately (excluding community and clinic-based programs) are crucial elements in the assessment of sex education effectiveness. This is especially true if that assessment intends to be of practical use to stakeholders such as school administrators, parents, and policymakers, in the effort to diminish the negative consequences of adolescent sexual activity.

Conclusions. ***Using meaningful and recommended criteria (12-month post-program effects for the intended population on key protective indicators), we found insufficient evidence of effectiveness by U.S. school-based CSE at decreasing teen pregnancy or STD rates or increasing teen abstinence or condom use. There was far more evidence of CSE failure than success. With regard to school-based AE, we found promising evidence for producing sustained increases in teen abstinence, enough to justify additional research.***

Recommendations. ***With regard to sex education in the United States: 1) Given the substantial evidence of program ineffectiveness as measured by meaningful standards from the field of prevention research, and the negative effects found in several studies, we do not recommend comprehensive sex education as a viable public health strategy in U.S. school classrooms.*** **2)** ***Replication studies should be conducted to verify the positive findings for school-based abstinence education, in order to better inform public policy.***

ENDNOTES

1. U.S. Centers for Disease Control and Prevention. (2016). *Sexually Transmitted Disease Surveillance 2015*. Retrieved from https://www.cdc.gov/std/stats15/STD-Surveillance-2015-print.pdf

2. United Nations Educational, Scientific and Cultural Organization. (2009). *International Technical Guidance on Sexuality Education, Volume 1*. Retrieved from http://unesdoc.unesco.org/images/0018/001832/183281e.pdf

3. Hamilton, B., Martin, J., Ventura, S. (2010). Births: Preliminary Data for 2009. *National Vital Statistics Reports 59*(3), 1–19; United Nations Statistics Division. (2008). *Demographic Yearbook 2006*. Retrieved from http://unstats.un.org/unsd/demographic/products/dyb/dyb2006.htm

4. Hallfors, D. D., Waller, M. W., & Ford, C. A., et al. (2004). Adolescent depression and suicide risk: association with sex and drug behaviors. *American Journal of Preventive Medicine 27*(3), 224–231; Sabia, J. J., Rees, D. I. (2008). The effect of adolescent virginity status on psychological well-being. *Journal of Health Economics, 27*(5), 1368–1381; Silverman, J. G., Raj, A., Clements, K. (2004). Dating violence and associated risk and pregnancy among adolescent girls in the United States. *Pediatrics, 114*(2), 220–225.

5. Long-acting reversible contraceptives (LARC) are methods of birth control that provide effective contraception for an extended period without requiring user action. They include injections, intrauterine devices (IUDs) and subdermal contraceptive implants. See https://en.wikipedia.org/wiki/Long-acting_reversible_contraception

6. For example, the *Planned Parenthood* organization is self-described as "the number one provider of sex education in the [United States]" including both "abstinence" and "safer sex" (another term for condom-based education) programming. (See https://www.plannedparenthood.org/get-care/our-services/patient-education.) Yet the "For Teens" section of the *Planned Parenthood* website contains the following messaging: "There's a lot to think about when it comes to sex: figuring out if you're ready, learning about orgasms, protecting yourself from pregnancy and STDs, how to know if someone wants to have sex with you, and much more … Stressing about whether you're a virgin is way less important than how you feel about your sexual experiences. Ask yourself: are you happy with the sexual experiences you've had or decided not to [have]? ... Sexually transmitted infections are super common—most people get one at some point in their life. Some STDs can be serious, but the good news is they can usually be cured or treated." (See https://www.plannedparenthood.org/learn/teens/sex; https://www.plannedparenthood.org/learn/teens/sex/virginity; https://www.plannedparenthood.org/learn/teens/preventing-pregnancy-stds.) These messages illustrate a lack of priority given to teen abstinence as a goal of prevention efforts, even though this organization purports to provide abstinence education.

7. United Nations Educational, Scientific and Cultural Organization. (2009). *International Technical Guidance on Sexuality Education, Volume 1*. Retrieved from http://unesdoc.unesco.org/images/0018/001832/183281e.pdf

8. Ibid.

9. See https://tppevidencereview.aspe.hhs.gov/EvidencePrograms.aspx

10. The development of standards for scientific evidence of program effectiveness has been undertaken by national entities like *The Society for Prevention Research (SPR), The What Works Clearinghouse*, and *Blueprints for Violence Prevention*. A consensus has been proposed by *SPR's* Standards of Evidence Committee in their publication, "Standards of Evidence: Criteria for Efficacy, Effectiveness, and Dissemination" (Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D. (2005). Standards of Evidence: Criteria for Efficacy, Effectiveness and Dissemination. *Prevention Science, 6*(3), 151–175), and recently updated (Gottfredson, D. C., Cook, T. D., Gardner, F. E. M., Gorman-Smith, D., Howe, G. W., et. al. (2015). Standards of Evidence for Efficacy, Effectiveness, and Scale-up Research in Prevention Science: Next Generation. *Prevention Science, 16* (7), 893–926. doi: 10.1007/s11121-015-0555-x. Retrieved from http://www.preventionresearch.org/wp-content/uploads/2011/12/Standards-of-Evidence_2015.pdf). These standards include a requirement of long-term sustained effects as well as a concern about main effects vs. subgroup effects.

11. See http://www.colorado.edu/cspv/blueprints/criteria.html

12. Gottfredson, D. C., Cook, T. D., Gardner, F. E. M., Gorman-Smith, D., Howe, G. W., et. al. (2015). Standards of Evidence for Efficacy, Effectiveness, and Scale-up Research in Prevention Science: Next Generation. *Prevention Science, 16* (7), 893–926. doi: 10.1007/s11121-015-0555-x. Retrieved from http://www.preventionresearch.org/wp-content/uploads/2011/12/Standards-of-Evidence_2015.pdf; Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D. (2005). Standards of Evidence: Criteria for Efficacy, Effectiveness and Dissemination. *Prevention Science, 6*(3), 151–175.

13. Ibid.

14. Ibid.

15. According to the CDC, "inconsistent use, e.g., failure to use condoms with every act of intercourse, can lead to STD transmission because transmission can occur with a single act of intercourse" (Centers for Disease Control and Prevention. (2003). *Fact Sheet for Public Health Personnel—Male Latex Condoms and Sexually Transmitted Diseases*. Retrieved from www.cdc.gov/nchstp/od/latex.htm). A study in the journal *AIDS* (*N*=17,264) found "Irregular condom use was not protective against HIV or STD and was associated with increased gonorrhea/Chlamydia risk" (Ahmed, S., Lutalo, T., Wawer, M., et al. (2001). HIV incidence and sexually transmitted disease prevalence associated with condom use: a population study in Rakai, Uganda. *AIDS, 15*(16), 2171–2179). A Denver study (*N*=26,291) reported that "Among the total population, rates of STD were higher among inconsistent [condom] users than nonusers … However, STD rates were significantly lower among consistent than inconsistent users" (Shlay, J. C., McCung, M. W., Patnaik, J. L., et al. (2004). Comparison of sexually transmitted disease prevalence by reported level of condom use among patients attending an urban sexually transmitted disease clinic. *Sexually Transmitted Diseases, 31*(3), 154–160). See also Crosby, R. A., DiClemente, R. J., Wingood, G. M., Lang, D., Harrington, K. F. (2003). Value of consistent condom use: A study of sexually transmitted disease prevention among African American adolescent females. *American Journal of Public Health; 93*(6), 901–902; and Grinsztejn, B., Veloso, V., Levi, J., Velasque, L., Luz, P., et al. (2009). Factors associated with increased prevalence of human papillomavirus infection in a cohort of HIV-infected Brazilian women. *International Journal of Infectious Diseases, 13*(1), 72–80.

16. Consistent condom use is the behavior upon which most estimates of the condom's protective capacity are based. The level of STD protection provided by consistent condom use ranges from a 30% risk reduction for genital herpes to 80% risk reduction for HIV transmission. See Martin, E. T., Krantz, E., Gottlieb, S. L., Magaret, A. S., Langenberg, A., et al. (2009). A Pooled Analysis of the Effect of Condoms in Preventing HSV-2 Acquisition. *Archives of Internal Medicine, 169*(13), 1233–1240; Weller, S. & Davis, K. (2002). Condom effectiveness in reducing heterosexual HIV transmission. *The Cochrane Database of Systemic Reviews, 1*; Sanchez, J., Campos, P., Courtois, B., Gutierrez, L., Carrillo, C., Alarcon, J., et al. (2003). Prevention of sexually transmitted diseases (STDs) in female sex workers: Prospective evaluation of condom promotion and strengthened STD services. *Sexually Transmitted Diseases*, *30*(4), 273–279; Holmes, K. K., Levine, R., Weaver, M. (2004). Effectiveness of condoms in preventing sexually transmitted infections. *Bulletin of the World Health Organization, 82*(6), 454–461.

17. Hallfors, D. D., Waller, M. W., & Ford, C. A., et al. (2004). Adolescent depression and suicide risk: association with sex and drug behaviors. *American Journal of Preventive Medicine 27*(3), 224–231; Sabia, J. J. & Rees, D. I. (2008). The effect of adolescent virginity status on psychological well-being. *Journal of Health Economics, 27,* 1368–1381; Silverman, J. G., Raj, A., Clements, K. (2004). Dating violence and associated risk and pregnancy among adolescent girls in the United States. *Pediatrics, 114*(2), 220–225.

18. A recent meta-analysis involving 12 studies in Sub-Saharan Africa concluded that women taking Depo-Provera (DMPA) had a somewhat elevated risk of contracting HIV (Ralph, L. J., et al. (2015). Hormonal contraceptive use and women's risk of HIV acquisition: a meta-analysis of observational studies. *The Lancet Infectious Diseases, 15*, 181–189). Another study found that use of DMPA more than doubled the risk of developing breast cancer in recipients (Li, C. I., et al. (2012). Effect of Depo-Medroxyprogesterone Acetate on Breast Cancer Risk among Women 20 to 44 Years of Age. *Cancer Research, 72*, 2028–2035. Retrieved from http://cancerres.aacrjournals.org/content/72/8/2028.

19. An interval that has been frequently used by researchers evaluating youth programs is 12 months or one year after the program. For example, "sustained impact," defined as "at least one year beyond treatment," is required by the "Blueprints Programs" of the Center for the Study and Prevention of Violence in order for an intervention to be designated as an effective or model program (see http://www.colorado.edu/cspv/blueprints/criteria.html), and long-term impact is defined by the federal 2010 *Teenage Pregnancy Prevention* initiative as an effect that is sustained for at least one year *after* program participation (see Office of Adolescent Health. (2010). *Teenage Pregnancy Prevention: Replication of Evidence-based Programs (Tier 1)—Funding Opportunity Announcement and Application Instructions*. Office of Public Health & Science, U.S. Department of Health and Human Services). The *Society for Prevention Research* cites a six-month follow-up as the *minimum* to demonstrate that program effects "do not dissipate immediately" and suggests a longer time-frame with multiple intervals during adolescence to assess behavioral effects in a teen population. (Gottfredson, D. C., Cook, T. D., Gardner, F. E. M., Gorman-Smith, D., Howe, G. W., et. al. (2015). Standards of Evidence for Efficacy, Effectiveness, and Scale-up Research in Prevention Science: Next Generation. *Prevention Science, 16* (7), 893–926. doi: 10.1007/s11121-015-0555-x. Retrieved from http://www.preventionresearch.org/wp-content/uploads/2011/12/Standards-of-Evidence_2015.pdf)

20. Gottfredson, D. C., Cook, T. D., Gardner, F. E. M., Gorman-Smith, D., Howe, G. W., et. al. (2015). Standards of Evidence for Efficacy, Effectiveness, and Scale-up Research in Prevention Science: Next Generation. *Prevention Science, 16*(7), 893-926. doi: 10.1007/s11121-015-0555-x. Retrieved from http://www.preventionresearch.org/wp-content/uploads/2011/12/Standards-of-Evidence_2015.pdf

21. Ibid.

22. See "Teen Outreach Program (TOP)" in Farb, A. & Margolis, A. (2016). The Teen Pregnancy Prevention Program (2010–2015): Synthesis of Impact Findings. *American Journal of Public Health*, v. 106 (Suppl 1); Philliber, A. E., Philliber, S., & Brown, S. (2015). Evaluation of the Teen Outreach Program® in The Pacific Northwest. Available at https://tppevidencereview.aspe.hhs.gov/EvidencePrograms.aspx

23. Gottfredson, D. C., Cook, T. D., Gardner, F. E. M., Gorman-Smith, D., Howe, G. W., et. al. (2015). Standards of Evidence for Efficacy, Effectiveness, and Scale-up Research in Prevention Science: Next Generation. *Prevention Science*, *16*(7), 893–926. Available at http://www.preventionresearch.org/wp-content/uploads/2011/12/Standards-of-Evidence_2015.pdf

24. For example, *SPR* reports that "the past decade has … witnessed a disturbingly high rate of failures to replicate when independent evaluation teams conduct studies of prevention interventions" and that "effect sizes from trials conducted by program developers/creators were more than twice the size of effect sizes from trials conducted by others." See Gottfredson, D. C., Cook, T. D., Gardner, F. E. M., Gorman-Smith, D., Howe, G. W., et. al. (2015). Standards of Evidence for Efficacy, Effectiveness, and Scale-up Research in Prevention Science: Next Generation. *Prevention Science*, *16* (7), 893–926Available at http://www.preventionresearch.org/wp-content/uploads/2011/12/Standards-of-Evidence_2015.pdf

25. This concern was raised by the review team for the U.S. Department of Health and Human Services Teen Pregnancy Prevention (*TPP*) Program: "[a]ll but one of the [original] program models meeting the standards of research quality demonstrated evidence of effectiveness through a single study, often conducted by the developer of the program. The review team noted the lack of replication studies as a gap in the evidence base and called for subsequent, independent evaluations to determine the effectiveness of the programs" (Farb, A. & Margolis, A. (2016). The Teen Pregnancy Prevention Program (2010–2015): Synthesis of Impact Findings. *American Journal of Public Health*, v. 106 (Suppl 1)).

26. Gottfredson, D. C., Cook, T. D., Gardner, F. E. M., Gorman-Smith, D., Howe, G. W., et. al. (2015). Standards of Evidence for Efficacy, Effectiveness, and Scale-up Research in Prevention Science: Next Generation. *Prevention Science*, *16*(7), 893–926. doi: 10.1007/s11121-015-0555-x. Retrieved from http://www.preventionresearch.org/wp-content/uploads/2011/12/Standards-of-Evidence_2015.pdf

27. They produced at least one statistically significant effect, of any duration or for any subgroup, on any desired outcome. (Farb, A. & Margolis, A. (2016). The Teen Pregnancy Prevention Program (2010–2015): Synthesis of Impact Findings. *American Journal of Public Health*, v. 106 (Suppl 1).)

28. See Farb, A. & Margolis, A. (2016). The Teen Pregnancy Prevention Program (2010–2015): Synthesis of Impact Findings. *American Journal of Public Health*, v. 106 (Suppl 1); https://tppevidencereview.aspe.hhs.gov/pdfs/Summary_of_findings_2016-2017.pdf; https://www.hhs.gov/ash/oah/evaluation-and-research/grantee-led-evaluation/2010-2014-grantees/index.html; and summaries of each study available at: https://www.hhs.gov/ash/oah/sites/default/files/ash/oah/oah-initiatives/evaluation/grantee-led-evaluation/summary-ebps.pdf, and https://www.hhs.gov/ash/oah/sites/default/files/ash/oah/oah-initiatives/evaluation/grantee-led-evaluation/summary-researchdemonstration.pdf

29. The original list of 28 prevention programs with "evidence of effectiveness" has been expanded to 48 currently. See: https://tppevidencereview.aspe.hhs.gov/EvidencePrograms.aspx. A subset of these are school-based CSE programs.

30. The *Community Preventive Services Task Force* was established by the U.S. Department of Health and Human Services (DHHS) in 1996 and operates under its auspices with support from the Centers for Disease Control and Prevention (CDC). For a report of the study findings, see Chin H. B., Sipe, T. A., Elder, R., Mercer, S. L., Chattopadhyay, S., et al. (2012). The Effectiveness of Group-Based Comprehensive Risk Reduction and Abstinence Education Interventions to Prevent or Reduce the Risk of Adolescent Pregnancy, HIV, and STIs: Two Systematic Reviews for the Guide to Community Preventive Services. *American Journal of Preventive Medicine, 42*(3), 272–294; Weed, S. E. (2012). Sex Education Programs for Schools Still in Question: A Commentary on Meta-Analysis, *American Journal of Preventive Medicine, 42*(3), 313–315; Community Preventive Services Task Force. (2011). Recommendations for Group-Based Behavioral Interventions to Prevent Adolescent Pregnancy, Human Immunodeficiency Virus, and Other Sexually Transmitted Infections: Comprehensive Risk Reduction and Abstinence Education. *American Journal of Preventive Medicine, 42*(3), 304–307, see p.305.

31. United Nations Educational, Scientific and Cultural Organization. (2009). *International Technical Guidance on Sexuality Education, Volume 1*, see pp.15-17, available at http://unesdoc.unesco.org/images/0018/001832/183281e.pdf; United Nations Educational, Scientific and Cultural Organization. (2018). *International Technical Guidance on Sexuality Education: An Evidence-Informed Approach*. Revised Edition, see pp.28-29. Available at: http://unesdoc.unesco.org/images/0026/002607/260770e.pdf

32. See: https://tppevidencereview.aspe.hhs.gov/EvidencePrograms.aspx

33. Respectively: Tortolero, S., Markham, C., Fleslcher, M., Shegog, R., Addy, R., et al. (2010). It's Your Game: Keep It Real: Delaying Sexual Behavior with an Effective Middle School Program. *Journal of Adolescent Health*, *46*(2), 169–179; Howard, M., & McCabe, J. (1990). Helping teenagers postpone sexual involvement. *Family Planning Perspectives, 22* (1), 21–26; Reyna, V. F., & Mills, B. A. (2014). Theoretically Motivated Interventions for Reducing Sexual Risk Taking in Adolescence: A Randomized Controlled Experiment Applying Fuzzy-Trace Theory. *Journal of Experimental Psychology*. *General*, *143*(4), 1627–1648. (See results for the modified *RTR* intervention.)

34. Potter, S., Coyle, K., Glassman, J., Kershner, S., & Prince, M. (2016). It's Your Game …Keep It Real in South Carolina: A Group Randomized Trial Evaluating the Replication of an Evidence-Based Adolescent Pregnancy and Sexually Transmitted Infection Prevention Program. *American Journal of Public Health*, *106*(S1), S60–S69.

35. Grossman, J. M., Tracy, A. J., Charmaraman, L., Ceder, I., & Erkut, S. (2014). Protective Effects of Middle School Comprehensive Sex Education with Family Involvement. *Journal of School Health*, *84*(11), 739–747.

36. Ekstrand, M. L., Siegel, D. S., Nido, V., Faigeles, B., Cummings, G. A., Battle, R., et al. (1996). Peer-led AIDS prevention delays onset of sexual activity and changes peer norms among urban junior high school students. *Paper presented at XI International Conference on AIDS*. Vancouver, British Columbia, Canada.

37. Tortolero, S., Markham, C., Fleslcher, M., Shegog, R., Addy, R., et al. (2010). It's Your Game: Keep It Real: Delaying Sexual Behavior with an Effective Middle School Program. *Journal of Adolescent Health, 46*(2), 169–179; Villarruel, A. M., Jemmott, J. B., & Jemmott, L. S. (2006). A randomized controlled trial testing an HIV prevention intervention for Latino youth. *Archives of Pediatrics & Adolescent Medicine*, 160(8), 772–777; Moberg, D., Piper, D. L. (1998). The Healthy for Life Project: Sexual risk behavior outcomes. *AIDS Education and Prevention*, *10*(2), 128–148.

38. Jemmott, J. B., III, Jemmott, L. S., & Fong, G. T. (2010). Efficacy of a theory-based abstinence-only intervention over 24 months: A randomized controlled trial with young adolescents. *Archives of Pediatrics & Adolescent Medicine, 164*(2), 152–159; Villarruel, A. M., Jemmott, J. B., & Jemmott, L. S. (2006). A randomized controlled trial testing an HIV prevention intervention for Latino youth. *Archives of Pediatrics & Adolescent Medicine*, *160*(8), 772–777; Reyna, V. F., & Mills, B. A. (2014). Theoretically Motivated Interventions for Reducing Sexual Risk Taking in Adolescence: A Randomized Controlled Experiment Applying Fuzzy-Trace Theory. *Journal of Experimental Psychology*. *General*, *143*(4), 1627–1648. (See results for the modified *RTR* intervention.)

39. According to the CDC, "inconsistent use (e.g., failure to use condoms with every act of intercourse), can lead to STD transmission because transmission can occur with a single act of intercourse" (Centers for Disease Control and Prevention. (2003). *Fact Sheet for Public Health Personnel—Male Latex Condoms and Sexually Transmitted Diseases*. Retrieved from www.cdc.gov/nchstp/od/latex.htm); see also Crosby, R. A., DiClemente, R. J., Wingood, G. M., Lang, D., Harrington, K. F. (2003). Value of consistent condom use: A study of sexually transmitted disease prevention among African American adolescent females. *American Journal of Public Health; 93*(6), 901–902.

40. Villarruel, A. M., Jemmott, J. B., & Jemmott, L. S. (2006). A randomized controlled trial testing an HIV prevention intervention for Latino youth. *Archives of Pediatrics & Adolescent Medicine*, *160*(8), 772–777.

41. Kelsey, M., Layzer, C., Layzer, J., Price, C., Juras, R., et. al. (2016). Replicating ¡Cuídate!: 6-Month Impact Findings of a Randomized Controlled Trial. *American Journal of Public Health, 106*(S1), S70–S77; Abt Associates. ¡Cuídate!: Interim Impact Report, Teen Pregnancy Prevention Replication Study, Report prepared for the Office of Adolescent Health and the Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services, September 2015.

42. Fisher, J., Fisher, W., Bryan, A., & Misovich, S. (2002). Information-motivation-behavioural skills model-based HIV risk behaviour change intervention for inner-city high school youth. *Health Psychology, 21*(2), 177-186; Coyle, K. K., Basen-Enquist, K. M., Kirby, D. B., Parcel, G. S., Banspach, S. W., Collins, J. L., et al. (2001). Safer Choices: Reducing Teen Pregnancy, HIV and STDs. *Public Health Reports, 1*(16), 82–93; Jemmott, J. B. III, Jemmott, L. S., Fong, G. T. (1998). Abstinence and safer sex HIV risk reduction interventions for African American adolescents. *Journal of American Medical Association, 279*(19), 1529–1536.

43. Respectively, Villarruel, A. M., Jemmott, J. B., & Jemmott, L. S. (2006). A randomized controlled trial testing an HIV prevention intervention for Latino youth. *Archives of Pediatrics & Adolescent Medicine*, 160(8), 772–777; Kelsey, M., Layzer, C., Layzer, J., Price, C., Juras, R., et. al. (2016). Replicating ¡Cuídate!: 6-Month Impact Findings of a Randomized Controlled Trial. *American Journal of Public Health, 106*(S1), S70–S77; Abt Associates. ¡Cuídate!: Interim Impact Report, Teen Pregnancy Prevention Replication Study, Report prepared for the Office of Adolescent Health and the Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services, September 2015.

44. Markham, C. M., Tortolero, S. R., Peskin, M. F., Shegog, R., Thiel, M., Baumler, E. R., Addy, R. C., Escobar-Chaves, S. L., Reininger, B., & Robin, L. (2012). Sexual risk avoidance and sexual risk reduction interventions for middle school youth: A randomized controlled trial. *Journal of Adolescent Health, 50*(3), 279–288; Markham, C. M., Peskin, M. F., Shegog, R., Baumler, E. R., Addy, R. C., Thiel, M., Escobar-Chaves, S. L.,

Robin, L., & Tortolero, S. R. (2014). Behavioral and psychosocial effects of two middle school sexual health education programs at tenth-grade follow-up. *Journal of Adolescent Health, 54*(2), 151–159.

45. Potter, S., Coyle, K., Glassman, J., Kershner, S., & Prince, M. (2016). It's Your Game …Keep It Real in South Carolina: A Group Randomized Trial Evaluating the Replication of an Evidence-Based Adolescent Pregnancy and Sexually Transmitted Infection Prevention Program. *American Journal of Public Health*, *106*(S1), S60–S69.

46. Respectively, Daley, E. M., Buhi, E. R., Wang, W., Singleton, A., Debate, R., Marhefka, S., et al. (2015). Evaluation of Wyman's Teen Outreach Program® in Florida: Final Impact Report for Florida Department of Health. Findings from the Replication of an Evidence-Based Teen Pregnancy Prevention Program; Allen, J. P., Philliber, S., Herrling, S., & Kuperminc, G. P. (1997). Preventing teen pregnancy and academic failure: Experimental evaluation of a developmentally based approach. *Child Development, 68*(4), 729–742; Philliber, A. E., Philliber, S., & Brown, S. (2015). Evaluation of the Teen Outreach Program® in The Pacific Northwest. Available at https://www.hhs.gov/ash/oah/sites/default/files/ash/oah/oah-initiatives/evaluation/grantee-led-evaluation/reports/ppgnw-final-report.pdf.; Francis, K., Philliber, S., Walsh-Buhi, E., Philliber, A., Seshadri, R., and Daley, E. (2016). Scalability of an Evidence-Based Adolescent Pregnancy Prevention Program: New Evidence From 5 Cluster-Randomized Evaluations of the Teen Outreach Program. *Am Journal of Public Health, 106,* S32–S38.

47. Villarruel, A. M., Jemmott, J. B., & Jemmott, L. S. (2006). A randomized controlled trial testing an HIV prevention intervention for Latino youth. *Archives of Pediatrics & Adolescent Medicine*, *160*(8), 772–777; Markham, C. M., Tortolero, S. R., Peskin, M. F., Shegog, R., Thiel, M., Baumler, E. R., Addy, R. C., Escobar-Chaves, S. L., Reininger, B., & Robin, L. (2012). Sexual risk avoidance and sexual risk reduction interventions for middle school youth: A randomized controlled trial. *Journal of Adolescent Health, 50*(3), 279–288; Main, D. S., Iverson, D. C., McGloin, J., Banspach, S. W., Collins, J., Rugg, D., et al. (1994). Preventing HIV infection among adolescents: Evaluation of a school-based education programme. *Preventive Medicine*, *23*(4), 409–417; Walter, H J. , & Vaughan, R. D. (1993). AIDS risk reduction among a multi-ethnic sample of urban high school students. *Journal of the American Medical Association, 270*(6), 725–730; Coyle, K. K., Kirby, D. B., Robin, L. E., Banspach, S. W., Baumler, E. R., Glassman, J. R. (2006). All4You! A randomized trial of an HIV, other STDs and pregnancy prevention intervention for alternative school students. *AIDS Education and Prevention, 18*(3), 187–203.

48. Reyna, V. F., & Mills, B. A. (2014). Theoretically Motivated Interventions for Reducing Sexual Risk Taking in Adolescence: A Randomized Controlled Experiment Applying Fuzzy-Trace Theory. *Journal of Experimental Psychology*. *General*, *143*(4), 1627–1648. (See results for the modified *RTR* intervention.)

49. Kirby, D., Barth, R. P., Leland, N., & Fetro, J. V. (1991). Reducing the Risk: Impact of a new curriculum on sexual risk-taking. *Family Planning Perspectives*, *23*(6), 253–263.

50. See: https://tppevidencereview.aspe.hhs.gov/document.aspx?rid=3&sid=182&mid=7

51. Hubbard, B. M., Giese, M. L., & Rainey, J. (1998). A replication of Reducing the Risk, a theory-based sexuality curriculum for adolescents. *Journal of School Health, 68* (6), 243–247; Also see: https://tppevidencereview.aspe.hhs.gov/document.aspx?rid=3&sid=182&mid

52. Zimmerman, R. S., Cupp, P. K., Donohew, L., Sionean, C. K., Feist-Price, S., & Helme, D. (2008). Effects of a school-based, theory-driven HIV and pregnancy prevention curriculum. *Perspectives on Sexual and Reproductive Health, 40*(1), 42–51. (See results for the two separate trials of both versions of *RTR*.)

53. See: https://tppevidencereview.aspe.hhs.gov/StudyDetails.aspx?id=21; and https://tppevidencereview.aspe.hhs.gov/document.aspx?rid=3&sid=182&mid=7#outcomes

54. Kelsey, M., Blocklin, M., Price, C., Juras, R., Freiman, L., et al. (2016). Replicating Reducing the Risk: 12-Month Impacts of a Cluster Randomized Controlled Trial. *American Journal of Public Health, 106*(S1), S45–S52; Reyna, V. F., & Mills, B. A. (2014). Theoretically Motivated Interventions for Reducing Sexual Risk Taking in Adolescence: A Randomized Controlled Experiment Applying Fuzzy-Trace Theory. *Journal of Experimental Psychology*. *General*, *143*(4), 1627–1648. (Within this article, see the evaluation of the standard version of RTR); Zimmerman, R. S., Cupp, P. K., Donohew, L., Sionean, C. K., Feist-Price, S., & Helme, D. (2008). Effects of a school-based, theory-driven HIV and pregnancy prevention curriculum. *Perspectives on Sexual and Reproductive Health, 40*(1), 42–51. (See results for the two separate trials of both versions of *RTR*.) Note: the Zimmerman study reported no effects for either of the two versions of *RTR* tested by the study, but reported a significant program effect on sexual initiation when the samples of the two different programs were combined. However, this "combined" effect appears to be an artifact since it did not occur in the real world—no adolescent received both versions of *RTR*. Moreover, since the programs were different enough to test against each other—apples and oranges—it does not seem appropriate to combine them and count this as evidence of an *RTR* effect.) Also of note: Another study of *RTR*—one of the eight mentioned here—by program developers, found a short-term six-month increase in teen contraceptive use, a less-protective indicator. See Barth, R. P. (1992). Preventing adolescent pregnancy with social and cognitive skills. *Journal of Adolescent Research*, *7*(2), 208-232.

55. Tortolero, S., Markham, C., Fleslcher, M., Shegog, R., Addy, R., et al. (2010). It's Your Game: Keep It Real: Delaying Sexual Behavior with an Effective Middle School Program. *Journal of Adolescent Health, 46*(2), 169–179.

56. Markham, C. M., Tortolero, S. R., Peskin, M. F., Shegog, R., Thiel, M., Baumler, E. R., Addy, R. C., Escobar-Chaves, S. L., Reininger, B., & Robin, L. (2012). Sexual risk avoidance and sexual risk reduction interventions for middle school youth: A randomized controlled trial. *Journal of Adolescent Health, 50*(3), 279–288; Markham, C. M., Peskin, M. F., Shegog, R., Baumler, E. R., Addy, R. C., Thiel, M., Escobar-Chaves, S. L., Robin, L., & Tortolero, S. R. (2014). Behavioral and psychosocial effects of two middle school sexual health education programs at tenth-grade follow-up. *Journal of Adolescent Health, 54*(2), 151–159. It should be noted that in both of these studies, the study write-up suggests that teens practicing abstinence were counted in the measure of consistent condom use by sexually active teens, conflating the two behaviors.

57. Potter, S., Coyle, K., Glassman, J., Kershner, S., & Prince, M. (2016). It's Your Game …Keep It Real in South Carolina: A Group Randomized Trial Evaluating the Replication of an Evidence-Based Adolescent Pregnancy and Sexually Transmitted Infection Prevention Program. *American Journal of Public Health*, *106*(S1), S60–S69.

58. Coyle, K., Anderson, P., Laris, B. A., Unti, T., Franks, H., & Glassman, J. (2015). Evaluation of It's Your Game: Keep It Real in Houston, TX: Final report.

59. Villarruel, A. M., Jemmott, J. B., & Jemmott, L. S. (2006). A randomized controlled trial testing an HIV prevention intervention for Latino youth. *Archives of Pediatrics & Adolescent Medicine*, *160*(8), 772–777.

60. Kelsey, M., Layzer, C., Layzer, J., Price, C., Juras, R., et. al. (2016). Replicating ¡Cuídate!: 6-Month Impact Findings of a Randomized Controlled Trial. *American Journal of Public Health, 106*(S1), S70–S77; Abt Associates. ¡Cuídate!: Interim Impact Report, Teen Pregnancy Prevention Replication Study, Report prepared for the Office of Adolescent Health and the Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services, September 2015.

61. Allen, J. P., Philliber, S., Herrling, S., & Kuperminc, G. P. (1997). Preventing teen pregnancy and academic failure: Experimental evaluation of a developmentally based approach. *Child Development, 68*(4), 729–742.

62. Daley, E. M., Buhi, E. R., Wang, W., Singleton, A., Debate, R., Marhefka, S., et al. (2015). Evaluation of Wyman's Teen Outreach Program® in Florida: Final Impact Report for Florida Department of Health. Findings from the Replication of an Evidence-Based Teen Pregnancy Prevention Program.

63. Francis, K., Woodford, M., and Kelsey, M. (2015). Evaluation of the Teen Outreach Program in Hennepin County, MN: Findings from the Replication of an Evidence-Based Teen Pregnancy Prevention Program.

64. Seshadri, R., Smithgall, C., Goerge, R., Ippolito, J., Dasgupta, D., Wiegand, E., Guiltinan, S., & Wood, M. (2015). Evaluation of Teen Outreach Program in Chicago: Final Impact Report for Chicago Public Schools.

65. Philliber, A. E., Philliber, S., & Brown, S. (2015). Evaluation of the Teen Outreach Program® in The Pacific Northwest.

66. Respectively, Howard, M., & McCabe, J. (1990). Helping teenagers postpone sexual involvement. *Family Planning Perspectives*, *22*(1), 21–26; Kirby, D., & Laris, B. A. (2007). Summaries of the Evaluations Referenced in Emerging Answers.

67. Aarons, S. J., Jenkins, R. R., Raine, T. R., El-Khorazaty, M. N., Woodward, K. M., Williams, R. L., et al. (2000). Postponing sexual intercourse among urban junior high school students: A randomized controlled evaluation. *Journal of Adolescent Health, 27*(4), 236–247.

68. Little, C. B., & Rankin, A. (2001). *An evaluation of the Postponing Sexual Involvement curriculum among upstate New York eighth graders.* Unpublished manuscript, State University of New York, Cortland.

69. Jemmott, J. B., Jemmott, L. S., Fong, G. T., & McCaffree, K. (1999). Reducing HIV risk-associated sexual behavior among African American adolescents: Testing the generality of intervention effects. *American Journal of Community Psychology*, *27*(2), 161–187.

70. Borawski, E. A., Trapl, E. S., Adams-Tufts, K., Hayman, L. L., Goodwin, M. A., & Lovegreen, L. D. (2009). Taking be proud! be responsible! to the suburbs: A replication study. *Perspectives on Sexual and Reproductive Health, 41*(1), 12–22.

71. Slater, H. M., and Mitschke, D. B. (2015). Evaluation of the *Crossroads* Program in Arlington, TX: Findings from an Innovative Teen Pregnancy Prevention Program.

72. Philliber, S, Kaye, J.W., Herring, S., West, E. (2002). Preventing pregnancy and improving health care access among teenagers: an evaluation of the Children's Aid Society—Carrera Program. *Perspectives on Sexual and Reproductive Health, 34*, 244–251.

73. Kirby, D. (2009). Reducing pregnancy and health risk behaviours in teenagers: Intensive, multicomponent programmes are not always effective. *BMJ, 339*, b2054.

74. Tucker, T. (2015). Evaluation of the Carrera Program: Findings from the replication of an evidence-based teen pregnancy prevention program. Atlanta, GA: *Tressa Tucker & Associates, LLC*; Herrling, S. (2016). Evaluation of the Children's Aid Society (CAS)-Carrera Adolescent Pregnancy Prevention Program in Chicago, IL: Findings from the Replication of an Evidence-Based Teen Pregnancy Prevention Program. Accord, NY: *Philliber Research & Evaluation*.

75. See *¡Cuídate!*, *It's Your Game: Keep It Real*, and *Teen Outreach Program*, under "Select a Program" at https://tppevidencereview.aspe.hhs.gov/EvidencePrograms.aspx; Gottfredson, D. C., Cook, T. D., Gardner, F. E. M., Gorman-Smith, D., Howe, G. W., et. al. (2015). Standards of Evidence for Efficacy, Effectiveness, and Scale-up Research in Prevention Science: Next Generation. *Prevention Science, 16*(7), 893–926. doi: 10.1007/s11121-015-0555-x. Retrieved from http://www.preventionresearch.org/wp-content/uploads/2011/12/Standards-of-Evidence_2015.pdf

76. Kelsey, M., Layzer, C., Layzer, J., Price, C., Juras, R., et. al. (2016). Replicating ¡Cuídate!: 6-Month Impact Findings of a Randomized Controlled Trial. *American Journal of Public Health, 106*(S1), S70–S77

77. Potter, S., Coyle, K., Glassman, J., Kershner, S., & Prince, M. (2016). It's Your Game …Keep It Real in South Carolina: A Group Randomized Trial Evaluating the Replication of an Evidence-Based Adolescent Pregnancy and Sexually Transmitted Infection Prevention Program. *American Journal of Public Health*, *106*(S1), S60–S69.

78. Philliber, A. E., Philliber, S., & Brown, S. (2015). Evaluation of the Teen Outreach Program® in The Pacific Northwest. Accord, NY: *Philliber Research & Evaluation*.

79. Moberg, D., Piper, D. L. (1998). The Healthy for Life Project: Sexual risk behavior outcomes. *AIDS Education and Prevention 10*(2):128–48.

80. Kirby, D., Korpi, M., Adivi, C., Weissman, J. (1997). An impact evaluation of SNAPP, a pregnancy-and-AIDS-prevention middle school curriculum. *AIDS Prevention and Education*, *9*,(1 Suppl), 44–61.

81. One major source of the perception that abstinence education is ineffective comes from the findings of six problematic AE studies: four produced in a 2007 evaluation by Mathematica Policy Research, Inc. (Trenholm, C., Devaney, B., Fortson, K., Quay, L., Wheeler, J., & Clark, M. (2007). *Impacts of four Title V,Section 510 abstinence education programs*. Princeton, NJ: Mathematica Policy Research) and two other studies erroneously treated as evaluations of AE (Clark, M. A., Trenholm, C., Devaney, B., Wheeler, J., & Quay, L. (2007). *Impacts of the Heritage Keepers® Life Skills Education component*. Princeton, NJ: Mathematica Policy Research, Inc.; Blake, S. M., Simkin, L., Ledsky, R., Perkins, C., & Calabrese, J. M. (2001). Effects of a Parent-Child Communications Intervention on Young Adolescents' Risk for Early Onset of Sexual Intercourse. *Family Planning Perspectives, 33*(2),52-61). These six studies have been cited by numerous reviewers as compelling evidence for AE failure. However, their design limitations raise concerns. For the Mathematica studies: 1) While touted as having a strong experimental (randomized) evaluation design, this methodology was weakened by randomizing the treatment and control groups *within* the same schools, disregarding the fact that cross contamination would likely occur between these two groups of youth—in the lunchroom, the locker room, and after-school programs, and within peer groups outside the school setting. Students tend to ignore their random group assignment and freely "share the medicine." And if the abstinence program reduces sexual behavior in the treatment group, it will also likely diminish this in the control group by reducing the number of sexual partners available to them. Thus, a reduction in sexual activity likely occurs in both groups as a result of the program, minimizing between group differences and the measurement of a program effect. 2) This design problem was compounded in the four studies by another methodological issue—the very young age of the program participants (ages 10-11, 11-13, 8-13, and 13). Measuring sexual behavior in a population this young typically finds such low rates that cell sizes are too small to produce statistically significant differences between program and control groups, even a year later. This limitation might have been addressed by employing appropriately longer follow-up time periods. Instead, a third major shortcoming occurred: 3) The follow-up time frames were so long—three to five years *after* the program (four to six years post baseline) and without any additional program message reinforcement during the interim—that a post-program effect on behavior could not have reasonably been expected to persist at that point. Such unusually long follow-up times have not been employed in CSE studies. These three factors in combination—randomizing within schools, unusually young subject populations, and unrealistically long follow-up time frames—argue for viewing the findings of these four studies as "inconclusive" rather than as valid evidence of AE program failure. For the Clark and Blake studies: Each of these measured the *additive effect* of a secondary program component—one was a voluntary after-school "life skills" component (that did not have abstinence as its focus), and the other was a parent-communication component—compared to the impact of the program's mandatory AE classroom curriculum alone, which served as the counterfactual in the study. *In both cases, the AE curriculum was the control condition, and the study was an evaluation of the impact of the subsidiary program component, not of the AE program. Yet these two studies have been treated as evaluations of AE classroom curricula in several important evidence reviews*. (For example, they, along with the four Mathematica studies, were included in the CDC-sponsored 2012 meta-analysis.) None of the six studies mentioned here found significant program effects, so their null findings combine to form a faulty evidence base that undermines the case for AE efficacy.

82. Chin H. B., Sipe, T. A., Elder, R., Mercer, S. L., Chattopadhyay, S., et al. (2012). The Effectiveness of Group-Based Comprehensive Risk Reduction and Abstinence Education Interventions to Prevent or Reduce the Risk of Adolescent Pregnancy, HIV, and STIs: Two Systematic Reviews for the Guide to Community Preventive Services. *American Journal of Preventive Medicine*, *42*(3), 272–294; Manlove, J., Fish, H. and Moore, K. A. (2015). Programs to improve adolescent sexual and reproductive health in the US: A review of the evidence. *Adolescent Health, Medicine and Therapeutics*, *6*, 47-79.

83. Jemmott, J. B., III, Jemmott, L. S., & Fong, G. T. (2010). Efficacy of a theory-based abstinence-only intervention over 24 months: A randomized controlled trial with young adolescents. *Archives of Pediatrics & Adolescent Medicine, 164*(2), 152–159; Erkut, S., Grossman, J. M., Frye, A. A., Ceder, I., Charmaraman, L., & Tracy, A. J. (2013). Can sex education delay early sexual debut? *Journal of Early Adolescence, 33*(4), 482–497 (Note: This was an abstinence curriculum that comprised the first year of a three-year program that was a CSE curriculum for the remaining two years. See: https://tppevidencereview.aspe.hhs.gov/document.aspx?rid=3&sid=274&mid=2); Weed, S. E., Ericksen, I. H., & Birch, P. J. (2005). An evaluation of the Heritage Keepers abstinence education program. In *Evaluating abstinence education programs: Improving implementation and assessing impact* (pp. 88–103).Washington, DC: Office of Population Affairs and the Administration for Children and Families, Department of Health & Human Services; Weed, S. E., Birch, P. J., Ericksen, I. H., & Olsen, J. A. (2011). Testing a predictive model of youth sexual intercourse initiation. Unpublished manuscript; Denny, G., & Young, M. (2006). An evaluation of an abstinence-only sex education curriculum: An 18-month follow-up. *Journal of School Health, 76*(8), 414–422; Weed, S. E., Ericksen, I. E., Lewis, A., et al. (2008). An Abstinence Program's Impact on Cognitive Mediators and Sexual Initiation. *American Journal of Health Behavior, 32*(1), 60–73; Weed, S. E., Anderson, N. A., Ericksen, I. E. (2008). What kind of abstinence education works? Comparing outcomes of two approaches. Salt Lake City: *Institute for Research & Evaluation*; Piotrowski, Z., Hedeker, H., & Hedeker, D. (2015). Evaluation of The Positive Potential Be The Exception Grade 6 Program in Predominantly Rural Communities: Findings from an Innovative Teen Pregnancy Prevention Program. Report to the Office of Adolescent Health, U.S. Department of Health & Human Services.

84. Walker, E. M., Inoa, R., & Coppola, N. (2016). Evaluation of Promoting Health Among Teens Abstinence-Only Intervention in Yonkers, NY; Farb, A. & Margolis, A. (2016). The Teen Pregnancy Prevention Program (2010–2015): Synthesis of Impact Findings. *American Journal of Public Health*, v. 106 (Suppl 1); Lieberman, L., & Su, H. (2012). Impact of the Choosing the Best Program in Communities Committed to Abstinence Education. *SAGE Open*, *2*(1), 1–12.

85. Borawski, E. A., Trapl, E. S., Lovegreen, L. D., Colabianchi, N., & Block, T. (2005). Effectiveness of abstinence-only intervention in middle school teens. *American Journal of Health Behavior, 29*(5), 423–434; Jemmott, J. B., III, Jemmott, L. S., & Fong, G. T. (2010). Efficacy of a theory-based abstinence-only intervention over 24 months: A randomized controlled trial with young adolescents. *Archives of Pediatrics & Adolescent Medicine, 164*(2), 152–159; Kirby, D., Korpi, M., Barth, R. P., & Cagampang, H. H. (1997). The impact of the Postponing Sexual Involvement curriculum among youths in California. *Family Planning Perspectives, 29*(3), 100–108; Trenholm, C., Devaney, B., Fortson, K., Quay, L., Wheeler, J., & Clark, M. (2007). Impacts of four Title V, Section 510 abstinence education programmes; Markham, C. M., Tortolero, S. R., Peskin, M. F., Shegog, R., Thiel, M., Baumler, E. R., Addy, R. C., Escobar-Chaves, S. L., Reininger, B., & Robin, L. (2012). Sexual risk avoidance and sexual risk reduction interventions for middle school youth: A randomized controlled trial. *Journal of Adolescent Health, 50*(3), 279–288; Markham, C. M., Peskin, M. F., Shegog, R., Baumler, E. R., Addy, R. C., Thiel, M., Escobar-Chaves, S. L., Robin, L., & Tortolero, S. R. (2014). Behavioral and psychosocial effects of two middle school sexual health education programs at tenth-grade follow-up. *The Journal of Adolescent Health, 54*(2), 151–159; Jemmott, J. B., Jemmott, L. S., & Fong, G. T. (1998). Abstinence and safer sex HIV risk-reduction interventions for African American adolescents: A randomized controlled trial. *Journal of the American Medical Association, 279*(19), 1529–1536. Piotrowski, Z., Hedeker, H., & Hedeker, D. (2015). Evaluation of The Positive Potential Be The Exception Grade 6 Program in Predominantly Rural Communities: Findings from an Innovative Teen Pregnancy Prevention Program. Report to the Office of Adolescent Health, U.S. Department of Health & Human Services; Piotrowski, Z., Hedeker, H., & Hedeker, D. (2015). Evaluation of The Positive Potential Be The Exception Grade 6 Program in Predominantly Rural Communities: Findings from an Innovative Teen Pregnancy Prevention Program. Report to the Office of Adolescent Health, U.S. Department of Health & Human Services.

86. Markham, C. M., Tortolero, S. R., Peskin, M. F., Shegog, R., Thiel, M., Baumler, E. R., Addy, R. C., Escobar-Chaves, S. L., Reininger, B., & Robin, L. (2012). Sexual risk avoidance and sexual risk reduction interventions for middle school youth: A randomized controlled trial. *Journal of Adolescent Health, 50*(3), 279–288; Markham, C. M., Peskin, M. F., Shegog, R., Baumler, E. R., Addy, R. C., Thiel, M., Escobar-Chaves, S. L., Robin, L., & Tortolero, S. R. (2014). Behavioral and psychosocial effects of two middle school sexual health education programs at tenth-grade follow-up. *Journal of Adolescent Health, 54*(2), 151–159.

87. See: https://tppevidencereview.aspe.hhs.gov/EvidencePrograms.aspx

88. Coyle, K. K., Basen-Enquist, K. M., Kirby, D. B., Parcel, G. S., Banspach, S. W., Collins, J. L., et al. (2001). Safer Choices: Reducing Teen Pregnancy, HIV and STDs. *Public Health Reports, 1*(16), 82–93; Jemmott, J. B., III, Jemmott, L. S., Fong, G. T. (1998). Abstinence and safer sex HIV risk reduction interventions for African American adolescents. *Journal of American Medical Association, 279*(19), 1529–1536.

89. Jemmott, J. B., III, Jemmott, L. S., & Fong, G. T. (2010). Efficacy of a theory-based abstinence-only intervention over 24 months: A randomized controlled trial with young adolescents. *Archives of Pediatrics & Adolescent Medicine, 164*(2), 152–159; Erkut, S., Grossman, J. M., Frye, A. A., Ceder, I., Charmaraman, L., & Tracy, A. J. (2013). Can sex education delay early sexual debut? *Journal of Early Adolescence, 33*(4), 482–497 (Note: This was an abstinence curriculum that comprised the first year of a three-year program that was a CSE curriculum for the remaining two years. See: https://tppevidencereview.aspe.hhs.gov/document.aspx?rid=3&sid=274&mid=2); Weed, S. E., Ericksen, I. H., & Birch, P. J. (2005). An evaluation of the Heritage Keepers abstinence education program. In *Evaluating abstinence education programs: Improving implementation and assessing impact* (pp. 88–103).Washington, DC: Office of Population Affairs and the Administration for Children and Families, Department of Health & Human Services; Weed, S. E., Birch, P. J., Ericksen, I. H., & Olsen, J. A. (2011). Testing a predictive model of youth sexual intercourse initiation. Unpublished manuscript; Piotrowski, Z., Hedeker, H., & Hedeker, D. (2015). Evaluation of The Positive Potential Be The Exception Grade 6 Program in Predominantly Rural Communities: Findings from an Innovative Teen Pregnancy Prevention Program. Report to the Office of Adolescent Health, U.S. Department of Health & Human Services.

90. Respectively: Jemmott, J. B., Jemmott, L. S., & Fong, G. T. (1998). Abstinence and safer sex HIV risk-reduction interventions for African American adolescents: A randomized controlled trial. *Journal of the American Medical Association, 279*(19), 1529–1536.
Jemmott, J. B., III, Jemmott, L. S., & Fong, G. T. (2010). Efficacy of a theory-based abstinence-only intervention over 24 months: A randomized controlled trial with young adolescents. *Archives of Pediatrics & Adolescent Medicine, 164*(2), 152–159; Piotrowski, Z., Hedeker, H., & Hedeker, D. (2015). Evaluation of The Positive Potential Be The Exception Grade 6 Program in Predominantly Rural Communities: Findings from an Innovative Teen Pregnancy Prevention Program. Report to the Office of Adolescent Health, U.S. Department of Health & Human Services.

91. United Nations Educational, Scientific and Cultural Organization. (2009). *International Technical Guidance on Sexuality Education, Volume 1*, p.15. Retrieved from http://unesdoc.unesco.org/images/0018/001832/183281e.pdf

92. Community Preventive Services Task Force. (2011). Recommendations for Group-Based Behavioral Interventions to Prevent Adolescent Pregnancy, Human Immunodeficiency Virus, and Other Sexually Transmitted Infections: Comprehensive Risk Reduction and Abstinence Education. *American Journal of Preventive Medicine, 42*(3), 304–307, see p.305; https://tppevidencereview.aspe.hhs.gov/EvidencePrograms.aspx.

93. See: http://www.advocatesforyouth.org/publications/1487

94. The methodological issues and weaknesses pertaining to these six studies are explained in detail in Endnote 81.

95. See: http://www.advocatesforyouth.org/publications/1487

96. Jemmott, J. B., III, Jemmott, L. S., & Fong, G. T. (2010). Efficacy of a theory-based abstinence-only intervention over 24 months: A randomized controlled trial with young adolescents. *Archives of Pediatrics & Adolescent Medicine, 164*(2), 152–159; Weed, S. E., Ericksen, I. H., & Birch, P. J. (2005). An evaluation of the Heritage Keepers abstinence education program. In *Evaluating abstinence education programs: Improving implementation and assessing impact* (pp. 88–103).Washington DC: Office of Population Affairs and the Administration for Children and Families, Department of Health & Human Services; Weed, S. E., Birch, P. J., Ericksen, I. H., & Olsen, J. A. (2011). Testing a predictive model of youth sexual intercourse initiation. Unpublished manuscript.

97. Community Preventive Services Task Force. (2011). Recommendations for Group-Based Behavioral Interventions to Prevent Adolescent Pregnancy, Human Immunodeficiency Virus, and Other Sexually Transmitted Infections: Comprehensive Risk Reduction and Abstinence Education. *American Journal of Preventive Medicine, 42*(3): 304–307, see p.305.

98. Weed, S. E. (2012). Sex Education Programs for Schools Still in Question: A Commentary on Meta-Analysis. *American Journal of Preventive Medicine, 42*(3), 313–315.

99. Derived from: The Community Preventive Services Task Force. (2008-2010). *Meta-Analysis Results: Condoms for CRR Interventions (Forest Plot)*. Data made available to Task Force members and consultants (including Irene H. Ericksen, co-author on the present article).

100. Denford, S., Abraham, C., Campbell, R., et al. (2017). A comprehensive review of reviews of school-based interventions to improve sexual-health. *Health Psychology Review*, *11*(1), 33–52., see pp.33, 39, 47.

101. See Benes, F. M. (1989). Myelination of cortical-hippocampal relays during late adolescence. *Schizophrenia Bulletin, 15*, 585–93; Benes, F. M. (1998). Brain development, VII. Human brain growth spans decades. *American Journal of Psychiatry, 155*, 1489; Frontline (Producer). (2002). *Inside the Teenage Brain: Interview with Dr. Deborah Yurgelun-Todd*. [Transcript from a television series episode]. Retrieved from http://www.pbs.org/wgbh/pages/frontline/shows/teenbrain/interviews/todd.html; Giedd, J., Blumenthal, J., Jeffries, N., Castellanos, F. X., Hong, L., Zijdenbos, A., et al. (1999). Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience, 2*, 861–863; Romanczyk, T. B., Weickert, C. S., Webster, M. J., Herman, M. M., & Kleinman, J. E. (2002). Alterations in the human prefrontal cortex across the life span. *European Journal of Neuroscience,15*, 269–280; Thompson, R. A. & Nelson, C. A. (2001). Developmental science and the media: Early brain development. *American Psychologist, 56*, 5–15; Yurgelun-Todd, D. (2002). *Frontline* interview Inside the Teenage Brain. Full interview available at http://www.pbs.org/wgbh/pages/frontline/shows/teenbrain/interviews/todd.html

102. Spinks, S. (2002). One reason teens respond differently to the world: Immature brain circuitry [Transcript from a television series episode]. In Frontline (Producer), *Inside the teenage brain*. Retrieved from http://www.pbs.org/wgbh/pages/frontline/shows/teenbrain/work/onereason.html, p. 2.

103. Thomas, M. (2000). Abstinence-based programs for prevention of adolescent pregnancies: A review. *Journal of Adolescent Health, 26*, 5–17.

104. Shlay, J. C., McClung, M. W., Patnaik, J. L., Douglas, J. M., Jr. (2004). Comparison of sexually transmitted disease prevalence by reported condom use: errors among consistent condom users seen at an urban sexually transmitted disease clinic. *Sexually Transmitted Diseases, 31*(9), 526–532; Grimley, D. M., Annang, L., Houser, S., Chen, H. (2004). Prevalence of Condom Use Errors Among STD Clinic Patients. *Sexually Transmitted Diseases, 31*(9), 526–532.

105. Sanders, S. A., Graham, C. A., Yarber, W. L., Crosby, R. A. (2003). Condom Use Errors and Problems Among Young Women Who Put Condoms on Their Male Partners, *Journal of the American Medical Women's Association, 58*(2), 95–98.

106. Tortolero, S. R., Markham, C. M., Shegog, R., Peskin, M. F. (2004). *It's Your Game: Keep it Real.* An HIV, STI, and Pregnancy, Curriculum for Middle Schools. *Center for Health Promotion and Prevention Research*, University of Texas-School of Public Health. Curriculum Manual, Level I Lessons, p.192, Level II Lessons.

107. Villarruel, A. M., Jemmott, L. S., Jemmott, J. B. (n.d.). *Facilitator's Curriculum, Module 2: Building Knowledge About Pregnancy, STDs and HIV*.

Table 1. U.S. School-based Comprehensive Sex Education (CSE): 60 Studies of 40 Programs*

PROGRAM & STUDY CHARACTERISTICS						PROGRAM OUTCOMES										
							Impact on Most-Protective Indicators for Intended Population				Less-Protective Indicators				Dual Benefit: Abstinence + Condom Use	
PROGRAM NAME	STUDY 1st AUTHOR & YEAR	Database	Independent Evaluator?	PROGRAM TYPE	Post-ProgramFollow-up Time (in Months)	Negative Effect	Sexual Initiation	Consistent Condom Use	Pregnancy	STDs	CondomFrequency	RecentSex	UnProtectedSex	# Sex Partners	12mo. After Program	Any Duration or Effect
1.Aban Aya (Curriculum Version Only)	Flay, 2004	TPP	No	CSE + Risk Behavior	<9(EndofAYearProg)	No	NM	NM	NM	NM	No	No	NM	NM	NM	No
2.AIDS Prevention Program	Siegel, 1995	CDC/UN	?	CSE	4	No	NM	NM	NM	NM	No	NM	NM	No	NM	No
3.All4You	Coyle, 2006	TPP/CDC/UN	No	CSE + Service-Learning	6, 12, 18	No	No	NM	No	NM	6 months Only	6 months Only	6 months Only	No	No	6 months
All4You	Coyle, 2013	TPP	No	CSE + Service-Learning	4, 16	No	No	NM	NM	NM	No	NM	No	NM	No	No
All4You (Curriculum Only)	Coyle, 2013	TPP	No	CSE Only	4, 16	No	No	NM	NM	NM	No	NM	4 months Only	NM	No	No
4.Be Proud Be Responsible (School-Recruited, held on Saturday)	Jemmott, 1999	TPP/CDC/UN	No	CSE	3, 6	No	NM	NM	NM	NM	NM	No	6 months	No	NM	NM
Be Proud Be Responsible (School Day/Classroom Version)	Borawski, 2009	TPP	Yes	CSE	4,12	No	No	No	NM	NM	NM	No	No	NM	NM	NM
5.(Blake) HIV/STD Prevention Curriculum	Blake, 2000 (unpub.)	UN	Yes	CSE	6	No	No	No	NM	NM	NM	No	NM	NM	NM	No
6.(Boyer) HIV/STD Prevention Curriculum	Boyer, 1997	CDC/UN	No	CSE	3	No	NM	NM	NM	NM	No	NM	NM	No	NM	No
7.Crossroads Program (adaptation BPBR)	Slater & Mitschke, 2015	TPP-T2	?	CSE	3, 6, 12	No	NM	NM	No	NM	NM	NM	6 months Only	NM	NM	NM
8.¡Cuídate! (School-Recruited, held on Saturday)	Villarruel, 2006	TPP/CDC/UN	No	CSE	12	No	No	12 months	NM	NM	No	12 months	12 months	12 months	No	12mo-RecentSx/#Partnrs
¡Cuídate! (School Day/Classroom Version)	Kelsey, 2016; Abt Assoc, 2015	TPP	Yes	CSE	6	Oral & Recent Sex	6mo(OralSex)Subgroup		NM	NM	NM	6mo-Subgroup	No	NM	NM	No
9.Draw the Line/Respect the Line	Coyle, 2004	TPP/CDC/UN	No	CSE	at Prog.End(3YrProg), 12	No	No		NM	NM	No	No	NM	No	No	No
10.Focus on Kids/West Virginia	Stanton, 2005	CDC/UN	No	CSE	3, 6, 9	No	NM	NM	NM	NM	No	No	NM	NM	NM	NM
11.Gender Matters	Smith,Kim,et al.2015 (Mathmj	TPP-T2	No	CSE	6	No	No	NM	NM	NM	NM	No	No	NM	NM	NM
12.Get Real About AIDS	Main, 1994	CDC/UN	?	CSE	6	No	No	NM	NM	NM	6 months	No	NM	6 months	NM	6months
13.Get Real - 7th & 8th Grade Only	Grossman, 2014	TPP	?	CSE	<9(End of 3 Year Prog)	No	<9 months	NM	NM	NM	NM	NM	NM	NM	NM	NM
14.H.A.R.T. (adaptation of B.A.R.T.)	Boston Medical Center	TPP-T2	?	CSE	6	No	No	NM	NM	NM	NM	No	No	NM	NM	No
15.Health Teacher	Mathmatica	TPP-T2	No	CSE	12	No	No	NM	NM	NM	NM	NM	NM	NM	NM	NM
16.Healthy & Alive!	Middlestadt, [Unpubl]	UN	No	CSE	<9	No	No	No	NM	NM	NM	No	NM	No	NM	No
17.Healthy for Life-Version 1 (Age-based)	Moberg, 1998/2000	CDC	?	CSE + Risk Behavior	24, 36, 48	No	No	No	NM	NM	NM	No	NM	NM	No	No
18.Healthy for Life-Version 2 (7th Gr.Intensive)	Moberg, 1998/2000	CDC	?	CSE + Risk Behavior	24, 36, 48	Recent Sex	No	No	NM	NM	NM	24 months	NM	NM	No	No
19.Healthy Oakland Teens	UN	?	?	CSE	8 to 11	No	8 to 11 months	NM	NM	NM	NM	NM	NM	NM	NM	NM
20.HIV Prevention Interventions	Fisher, 2002	CDC/UN	No	CSE	12	No	No	NM	NM	NM	12 months	NM	NM	NM	No	No
21.It's Your Game--Keep it Real	Tortolero, 2010	TPP	No	CSE	12	No	12 months	NM	NM	NM	No	12 months	NM	No	No	No
IYG - RR EOU/10moFup	Markham, 2012	TPP	No	CSE	10	No	10 months	10 months?**	NM	NM	NM	10 months	10 months	No	NM	10months
IYG - RR 24moFup	Markham, 2014	TPP	No	CSE	24	No	No-Anal Sex O	No-Anal Sex O?**	NM	NM	NM	No	No	No	No	24 mo-Anal Sex Only
IYG - RR	Potter, 2016 ("a" So.Car.)	TPP	Yes	CSE	12	Sexual Initiation	12 months		NM	NM	NM	No	No	NM	No	No
IYG - RR	Coyle, 2016 ("b" Texas)	TPP	Yes	CSE	12	No	No		NM	NM	NM	NM	NM	NM	NM	NM
22.Making Proud Choices! (School-Recruited, held on Saturday)	Jemmott, 1998	TPP/CDC/UN	No	CSE	3, 6, 12	No	No	3 months Only	NM	NM	3, 6, 12 months	No	No	NM	No	No
23.Need To Know	Jodi T.KHIthSciCtrSanAntonio	TPP-T2	?	CSE?	at Program End	No	No	NM	NM	NM	NM	NM	NM	NM	NM	No
24.Positive Prevention	LaChausse, 2006	CDC/UN	Yes	CSE	6	No	6 months	NM	NM	NM	No	6 months	NM	NM	NM	No
25.Positive Prevention PLUS	LaChausse, 2015/2016	TPP	Yes	CSE	6	No	6 months	NM	NM	No	NM	NM	6 months	NM	NM	NM
26.Postponing Sexual Involvement(PSI)	Howard&McCabe, 1990	UN	No	CSE	12	No	12 months	NM	NM	NM	NM	No	NM	NM	NM	NM
Postponing Sexual Involvement	Aarons, 2000	UN	Yes	CSE	6	No	No	NM	NM	NM	NM	NM	NM	NM	NM	NM
PSI & HumanSexuality-adapted	Little & Rankin, unpub.	UN	?	CSE	2, 6	No	No	NM	NM	NM	NM	No	NM	No	NM	NM
27.Project IMPPACT Inwood House	Lieberman, 2000	CDC/UN	No	CSE	at Prog.End, 12	No	No	NM	No	NM	No	NM	NM	NM	No	No
28.Project LIGHT	Lightfoot, 2007	CDC	No	CSE	3	No	NM	NM	NM	NM	NM	3 months	No	3 months	NM	No
29.Project SNAPP	Kirby et al., 1997	CDC/UN	Yes	CSE	5, 17	Contraceptive Use	No	NM	No	No	Contraception-17mo	No	NM	NM	No	No
30.Promoting Health Among Teens /CSE	Jemmott, 2010	TPP	No	CSE	24	No	No	No	NM	NM	NM	No	No	24 months	No	No
31.Reach for Health	O'Donnell, 1999	CDC	?	CSE (ClssrmCurrOnly)	6	No	No	NM	NM	NM	NM	No	NM	NM	NM	NM
32.Reducing the Risk (RTR)	Kirby et al., 1991	TPP/UN	No	CSE	18	No	No	NM	No	NM	No	NM	No	NM	NM	NM
RTR	Barth, 1992	TPP/CDC	No	CSE	6	No	No	NM	No	NM	No	No	NM	NM	NM	No
RTR	Hubbard, 1998	CDC/UN	Yes	CSE	18	No	18months	NM	NM	NM	NM	NM	NM	NM	NM	NM
RTR	Kelsey (Abt Assoc), 2016	TPP	Yes	CSE	12	No	No	NM	NM	NM	NM	No	No	NM	No	No
RTR	Zimmerman, 2008a	TPP/UN	No	CSE	>12	No	No	NM	NM	NM	No	NM	NM	NM	NM	No
33. RTR-modified1	Zimmerman, 2008b	TPP/UN	No	CSE	>12	No	No	NM	NM	NM	No	NM	NM	NM	NM	No
RTR	Reyna & Mills, 2014a	TPP	Yes	CSE	3, 6, 12	No	No	NM	NM	No	NM	No	No	No	No	No
34. RTR-modified2	Reyna & Mills, 2014b	TPP	No	CSE	3, 6, 12	No	12 months	NM	NM	NM	No	NM	No	12 months	No	No
35.Rochester AIDS Prevention Project(RAPP)	Siegel, 2001	CDC/UN	?	CSE	12	No	No	NM	NM	NM	NM	NM	NM	NM	NM	NM
Rochester AIDS Prevention Project(RAPP)	Aten, 2002	CDC	?	CSE	12	No	No	NM	NM	NM	NM	NM	NM	NM	NM	NM
36.Safer Choices	Coyle, 2001	TPP/CDC/UN	No	CSE	at Prog.End(2YrPrg), 12	No	No	NM	NM	NM	12 months	No	NM	No	No	No
37.Teen Outreach Program (TOP)	Allen1997/Philliber1992	TPP	Yes	Service-Learning+CSE	at ProgEnd(1YrPrg)	No	NM	NM	at Program End	NM	NM	NM	NM	NM	NM	NM
TOP	Oxley, 2015	TPP	Yes	Service-Learning+CSE	at ProgEnd(1YrPrg), 10	No	at Prog. End O	NM	at Program End O	NM	NM	NM	NM	NM	NM	NM
TOP	Francis (Hennepin,MN), 2015	TPP	Yes	Service-Learning+CSE	3, 15	No	No	NM	NM	NM	NM	No	No	NM	NM	NM
TOP	Seshadri,et al 2015 (Chicago)	TPP	Yes	Service-Learning+CSE	at Program End	No	NM	NM	NM	NM	NM	No	No	NM	NM	NM
TOP	Philliber, et al 2016 (GNYAPP)	TPP	Yes	Service-Learning+CSE	at Prog.End(1YrPrg), 12	Pregnancy-Females	NM	NM	At Prog. End (F)	NM	NM	No	NM	NM	NM	NM
38.[Walter&Vaughn]AIDS Prevention	Walter & Vaughn, 1993	CDC/UN	?	CSE	3	No	No	3 months	NM	No	NM	NM	NM	3 months	NM	3 months
39.Wise Guys	Sottsegen & Philliber 2001	UN	No	CSE	6	No	No	NM	NM	NM	NM	NM	NM	NM	NM	NM
40.Youth AIDS Prevention Project (YAPP)	Levy, 1995	CDC/UN	No	CSE	at Prog.End, 12	No	No	NM	NM	NM	No	No	NM	No	No	No

*Three governmental agencies (TPP, CDC, & UN--see Key below) conducted reviews of all credible studies of CSE in the US and claimed to have found evidence of CSE effectiveness. We examined the studies of *school-based* CSE included in these reviews & entered their results in this table.

**The study write-up did not make it clear whether only teens practicing CCU were counted here or whether teens practicing abstinence were also included in the measure, thus conflating the two behaviors.

KEY:

TPP/CDC/UN = U.S. Teen Pregnancy Prevention program/Centers for Disease Control & Prevention/UNESCO--These were the three government agencies whose databases we examined.

TPP-T2 = TPP's Tier Two funding of new "demonstration" programs/studies. Unlike the other TPP programs in Table 1, not all of these Tier Two programs met the TPP criteria for "evidence of effectiveness," but their studies constitute some of the most recent and rigo

NM = Did not measure this outcome; O = Only; F = Females; M = Males; at Prog. End= measured at end of program

No = The study measured this outcome but the effect was not statistically significant.

Green = Evidence of Success: A significant effect at least 12 months post-program, on the intended target population (a "main effect"), not just a subgroup.

Blue = Evidence of Potential: A significant main effect >3 months and <12 months post-program, or a subgroup effect >3 months, or a dual benefit <12 months or on less-protective indicators.

Tan - Evidence of Failure: measured short-term outcomes and failed to find significant effects.

Brown = Evidence of Failure: measured at least 12 months after the program and failed to find significant main effects.

Dark Brown = Evidence of Failure: measured a dual benefit for abstinence and condom outcomes for any time period but did not find significance for both.

Red = Evidence of Harm: A significant increase in some risk indicators for the main population or a substantial subgroup.

Light Grey = No evidence of negative impact.

Clear = Lack of Evidence: did not measure the outcome; or only measured a short-term effect (less than 6 months post-program) with significant results.

Yellow = Lack of Independent Evaluator: the study was conducted by the program's developer.

Light Purple = This program is on the TPP list of programs purportedly showing "evidence of effectiveness."

Table 2. Evidence of CSE Failure from 60 Studies of U.S. School-Based CSE Programs[a]

Measured Main Effects[b]	Reduced Teen Pregnancy		Reduced STDs		Abstinence (delayed sexual initiation)		Consistent Condom Use (CCU)		Any Condom Use (frequency, at last sex, CCU)		Reduction in Unprotected Sex (sex w/o condom or contraceptive)		Dual Benefit (Abstinence & Condom Use)	
Post-Program Duration[c]	At least 12 months	Any duration	At least 12 months	Any duration	At least 12 months	Any duration	At least 12 months	Any duration	At least 12 months	Any duration	At least 12 months	Any duration	At least 12 months	Any duration
Not Measured	54	50	59	58	28	10	54	51	43	32	34	37	40	25
Measured	6	10	1	2	32	50	6	9	17	28	16	23	20	36
POSITIVE EFFECT	0	2	0	0	4	10	1	3	4	7	1	8	0	6
NO EFFECT	6	7	1	2	27	38	5	6	13	21	15	15	20	30
NEGATIVE EFFECT	0	1	0	0	1	2	0	0	[1][d]	[1][d]	0	0	0	[3][e]
Failure Rate[f]	**LNS[g]**	**LNS[g]**	**LNS[g]**	**LNS[g]**	**88%**	**80%**	**LNS[g]**	**LNS[g]**	**76%**	**75%**	**94%**	**65%**	**100%**	**92%**

[a] Three governmental agencies (HHS/TPP, HHS/CDC, & UNESCO) conducted reviews of all credible studies of CSE in the US and claimed to have found evidence of CSE effectiveness. This table summarizes the evidence of program failure for the school-based CSE programs found in that database.

[b] A significant effect for the intended target population (i.e., a "main effect"), not just for a subgroup of the intended population of program participants.

[c] Shows two columns under each outcome: "At least 12 months" gives the findings, for each outcome, of studies measuring effects at least 12 months after the program—this is an important indicator of program effectiveness. "Any duration" gives findings of studies measuring effects of any duration or time period after the program, including immediate post-program effects, effects of short-term duration (3/6/9 months after the program) and long-term effects (at least 12 months after the program).

[d] One program produced a negative effect (reduction) in contraceptive use, which is not a specific measure of condom use but it is shown as a negative effect in this column since it may include condom use. However, it is not included in the column totals for that reason.

[e] Three of the 6 studies producing a dual benefit of any duration were representing 2 programs that were shown in other independent studies to have *reduced* teen abstinence—one of the dual benefits. For this reason those programs and their 3 studies are subtracted from the 6 showing a dual benefit in the net calculation, thus, 33/36 studies (92%) did not show effectiveness at a dual benefit of any duration.

[f] Of the studies in this database that measured an outcome (e.g., "Abstinence, at least 12 months after the program"), the proportion finding either a null effect (not statistically significant) or a statistically significant negative effect.

[g] Little or No Success (LNS): All or most effects found were null or negative, but the number of studies measuring this outcome was too small to estimate a numerical failure rate.

Table 3A. U.S. *Teen Pregnancy Prevention* (*TPP*) School-based Comprehensive Sex Education (CSE): 36 Studies of 18 Programs*

PROGRAM & STUDY CHARACTERISTICS					PROGRAM OUTCOMES											
PROGRAM NAME	STUDY 1st AUTHOR & YEAR	Independent Evaluator?	PROGRAM TYPE	Post-Program Follow-up Time (in Months)	Negative Effects	Impact on Most-Protective Indicators for Intended Population				Less-Protective Indicators				Dual Benefit: Abstinence+Condom Use	
						Sexual Initiation	Consistent Condom Use	Pregnancy	STDs	Any Condom Use	Recent Sex	Unprotected Sex	# of Sex Partners	12 mo. After Program	Any Duration or Effect
1.Aban Aya (Classroom Curriculum Component Only)	Flay, 2004	No	CSE + Risk Behavior	<9 (after 4 YearProg)	No	NM	NM	NM	NM	No	No	NM	NM	NM	No
2.All4You	Coyle, 2006	No	CSE + Service-learning	6, 12, 18	No	No	NM	No	NM	6 months Only	NM	NM	No	No	No
All4You	Coyle, 2013	No	CSE + Service-learning	4, 16	No	No	NM	NM	NM	No	NM	4 months Only	No	No	No
3.BeProudBeResponsible(School-recruited/After-school)	Jemmott, 1999	No	CSE	3, 6	No	NM	NM	NM	NM	NM	No(AnalSex O-6mo)	6 months (BC)	No (Anal Sex O-6mo)	NM	NM
4.BeProudBeResponsible(School Day/ClassroomVersion)	Borawski, 2009	Yes	CSE	4,12	No	No	No	NM	NM	NM	NM	No	NM	NM	NM
5.CAS Carrera(School-recruited/After-school+CommunityProg)	Philliber, 2002	Yes	Yth Dev,CSE&HlthClinic	at ProgEnd(3YearProg)	No	No (FO-at Prog.End)	NM	No (FO-at Prog.End)	NM	No	NM	NM	NM	NM	No
CAS Carrera	Kirby, 2005**	Yes	Yth Dev,CSE&HlthClinic	13,31(EndOfSYr.Prog)	Pregnancy	No	NM	During 3Yr Prog (FO)	NM	No	No	No	No	NM	No
[YPDP-CAS Carrera-model]	Wiggins, 2006**	Yes	Yth Dev,CSE&HlthClinic	at ProgEnd(1YrProg), 6	Sexual Initiation&Pregnancy	6 months (FO)	NM	6 months (FO)	NM	No	NM	NM	No	NM	No
CAS Carrera	Scher, 2009**	Yes	Yth Dev,CSE&HlthClinic	at Prog.End (3Yr.Prog)	No	NM	NM	NM	NM	NM	No	No	NM	NM	NM
CAS Carrera	Tucker, 2015	Yes	Yth Dev,CSE&HlthClinic	End OfYr1/2/3(3YrProg)	No	No(EndOfYr1, notYr2/3)	NM	NM	NM	NM	NM	No	NM	NM	NM
CAS Carrera	Herrling, 2016	Yes	Yth Dev,CSE&HlthClinic	at Prog.End (4YearProg)	No	No	NM	NM	NM	NM	NM	No	NM	NM	NM
6.iCuidate! (School-recruited/After-school Version)	Villaruel, 2006	No	CSE	12	No	No	12 months	NM	NM	No	12 months	12 months	12 months	No	12mo-RecentSx/#Partnrs
7.iCuidate! (School Day/Classroom Version)	Kelsey, 2016	Yes	CSE	6	Oral & Recent Sex	6mo(OralSex)Subgroup	NM	NM	NM	NM	6mo-Subgroup	No	NM	NM	No
8.Draw the Line/Respect the Line	Coyle, 2004	No	CSE	at Prog.End, 12	No	No (MO)	NM	NM	NM	No	No	NM	No	No	No
9.Get Real - 7th & 8th Grade Only	Grossman, 2014	?	CSE	<9	No	<9mo	NM	NM	NM	NM	NM	NM	NM	NM	NM
10.It's Your Game--Keep it Real	Tortolero, 2010	No	CSE	12	No	12 mo(Int/Oral&Anal)	NM	NM	NM	No	12 months	NM	No	No	No
IYG - RR EOU/10moFup	Markham, 2012	No	CSE	10	No	10 months	10 months****	NM	NM	NM	10 mo (Frequency)	10 months	No	NM	10months
IYG - RR 24moFup	Markham, 2014	No	CSE	24	No	No-Anal Sex O	No-Anal Sex O***	NM	NM	NM	No (Anal Sex O)	No (Anal Sex O)	No	No	24mo-Anal Sex Only
IYG - RR	Potter, 2016 ("a" So.Car.)	Yes	CSE	12	Sexual Initiation	12 months	NM	NM	NM	NM	No	No	NM	No	No
IYG - RR	Coyle, 2016 ("b" Texas)	Yes	CSE	12	No	No	NM	NM	NM	NM	NM	NM	NM	NM	NM
11.MakingProudChoices!(School-recruited/After-school)	Jemmott, 1998	No	CSE	3, 6, 12	No	No	3 months Only	NM	NM	3, 6, 12 months	No (SbGp6/12m)	3 mo(SbGp6/12m)	NM	No	No
12.Positive Prevention PLUS	LaChausse, 2015/2016	Yes	CSE	6	No	6 months	NM	No	NM	NM	NM	6 months (BC)	NM	NM	NM
13.PromotingHealthAmong Teens!(School-recruited/After-school)	Jemmott, 2010	No	CSE	24	No	No	No	NM	NM	NM	No	No	24 months	No	No
14.Reducing the Risk (RTR)	Kirby et al, 1991	No	CSE	18	No	No	NM	No	NM	NM	NM	No (BC)	NM	NM	NM
RTR	Barth, 1992	No	CSE	6	No	No	NM	No	NM	No (Subgroup O)	No	NM	NM	NM	No
RTR	Kelsey (Abt Assoc), 2016	Yes	CSE	12	No	No	NM	NM	NM	NM	No (Subgroup O)	No	NM	No	No
RTR (2ndEd.)	Zimmerman, 2008	No	CSE	>12	No	No	NM	NM	NM	No	NM	NM	NM	NM	No
15.Reducing the Risk-modified(A)	Zimmerman, 2008	No	CSE	>12	No	No	NM	NM	NM	No	NM	NM	NM	NM	No
RTR	Reyna & Mills, 2014	Yes	CSE	3, 6, 12	No	No	NM	NM	NM	No	NM	No	No	No	No
16.Reducing the Risk-modified(B)	Reyna & Mills, 2014	No	CSE	3, 6, 12	No	3, 6, 12 months	NM	NM	NM	No	NM	No	3, 6, 12 months	No	No
17.Safer Choices	Coyle, 2001	No	CSE	at Prog.End, 12	No	No	NM	NM	NM	12 months	NM	NM	No	No	No
18.Teen Outreach Program (TOP)	Allen1997/Philliber1992	Yes	Service-Learning+CSE	at Program End	No	NM	NM	NM	at Program End	NM	NM	NM	NM	NM	NM
TOP	Daley, 2015	Yes	Service-Learning+CSE	at Prog.End, 10	No	at Prog. End O	NM	NM	at Program End O	NM	NM	NM	NM	NM	NM
TOP	Francis (Abt&Assoc), 2015	Yes	Service-Learning+CSE	3, 15	No	No	NM	NM	NM	NM	No	No (BC)	NM	NM	NM
TOP	Seshadri, 2015 (Chicago)	Yes	Service-Learning+CSE	at Program End	No	NM	NM	NM	NM	NM	No	No	NM	NM	NM
TOP	Philliber, 2016 (GNWPP)	Yes	Service-Learning+CSE	at Prog.End, 12	Pregnancy	NM	NM	12 months (FO)	NM	NM	at Prog. End Only	NM	NM	NM	NM

* These programs are listed on the TPP website as having shown " evidence of effectiveness in reducing teen pregnancy, sexually transmitted infections, and associated sexual risk behaviors." See https://tppevidencereview.aspe.hhs.gov/EvidencePrograms.aspx

** These studies are not in the TPP database, but were reviewed in a commentary by Douglas Kirby, Ph.D., published in the British Medical Journal (BMJ, 2009), as representing the evidence base for this program. The Wiggins study was also published in the B

*** The study write-up did not make it clear whether only teens practicing CCU were counted here or whether teens practicing abstinence were also included in the measure, thus conflating the two behaviors.

KEY:

NM = Did not measure this outcome; O = Only; F = Females; M = Males; BC = Birth Control

No = Measured the outcome but the effect was not statistically significant.

Prog. End= Measured at End of Program.

Green = Evidence of Success: A significant effect at least 12 months post-program, on the intended target population (a "main effect"), not just a subgroup.

Blue = Evidence of Potential: A significant main effect >3 months and <12 months post-program, or a subgroup effect >3 months, or a dual benefit <12 months or on less-protective indicators.

Tan - Evidence of Failure: Measured short-term outcomes and failed to find significant effects.

Brown = Evidence of Failure: Measured at least 12 months after the program and failed to find significant main effects.

Dark Brown = Evidence of Failure: Measured a dual benefit for abstinence and condom outcomes for any time period but did not find significance for both.

Red = Evidence of Harm: A significant increase in some risk indicators for the main population or a substantial subgroup.

Light Grey = No evidence of negative impact.

Clear = Lack of Evidence: Did not measure the outcome; or only measured a short-term effect (less than 6 months post-program) with significant results.

Yellow = Lack of Independent Evaluator: The study was conducted by the program's developer.

Table 3B. U.S. *Teen Pregnancy Prevention* (*TPP*) School-based Abstinence Education (AE): 5 Studies of 5 Programs*

PROGRAM & STUDY CHARACTERISTICS					PROGRAM OUTCOMES										
						Impact on Most-Protective Indicators for Intended Population				Less-Protective Indicators				Dual Benefit: Abstinence+Condom Use	
PROGRAM NAME	STUDY 1st AUTHOR & YEAR	Independent Evaluator?	PROGRAM TYPE	Post-ProgramFollow-up in Months	Negative Effects	Sexual Initiation	Consistent Condom Use	Pregnancy	STDs	Any Condom Use	Recent Sex	UnprotectedSex	# of Sex Partners	12mo. After Program	Any Duration or Effect
Get Real-6th Grade Only **	Erkut, 2012	?	6thGrade Abstinence-only	12	No	12 months	NM	NM	NM	NM	NM	NM	NM	NM	NM
Heritage Keepers	Weed, 2011	Yes	Abstinence-only	12	No	12 months	NM	NM	NM	NM	NM	NM	NM	NM	NM
Making a Difference! An Abstinence Program	Jemmott, 1998	No	Abstinence-only	3, 6, 12	No	3 months Only	No	NM	NM	12 months	NM	No	NM	No	Abst(3m o)/Condoms(12mo)
Positive Potential - 6th Grade	Piotrowski, 2016	Yes	Youth Development + AE	3,12	No	12 months	No	NM	NM	NM	12 months	NM	NM	No	No
Promoting Health Among Teens!/Abstinence-Only	Jemmott, 2010	No	Abstinence-only	24	No	24months	No	NM	NM	NM	24months	No	No	No	No

* These programs are listed on the TPP website as having shown "evidence of effectiveness in reducing teen pregnancy, sexually transmitted infections, and associated sexual risk behaviors." See https://tppevidencereview.aspe.hhs.gov/EvidencePrograms.aspx

** This was a sex education program in which the 1st year (6th grade) component met the description of an abstinence-only curriculum (no contraception content) and in the 2nd & 3rd years (7th & 8th grade) it was comprehensive sex education (i.e., induded c

The 1st year component was evaluated separately and the results are reported here.

KEY:

NM = Did not measure this outcome; O = Only; F = Females; M = Males

No = Measured the outcome but the effect was not statistically significant.

Green = Evidence of Success: A significant effect at least 12 months post-program, on the intended target population (a "main effect"), not just a subgroup.

Blue = Evidence of Potential: A significant main effect >3 months and <12 months post-program, or a subgroup effect >3 months, or a dual benefit <12 months or on less-protective indicators.

Tan - Evidence of Failure: Measured short-term outcomes and failed to find significant effects.

Brown = Evidence of Failure: Measured at least 12 months after the program and failed to find significant main effects.

Dark Brown = Evidence of Failure: measured a dual benefit for abstinence & condom outcomes for any time period but did not find significance for both.

Red = Evidence of Harm: A significant increase in some risk indicators for the main population or a substantial subgroup.

Light Grey = No evidence of negative impact.

Clear = Lack of Evidence: Did not measure the outcome; or only measured a short-term effect (less than 6 months post-program) with significant results.

Yellow = Lack of Independent Evaluator: The study was conducted by the program's developer.