Re-Examining the Evidence for Comprehensive Sex Education in Schools

2019

A Global Research Review

Stan E. Weed, Ph.D.

Irene H. Ericksen, M.S.

The Institute for Research & Evaluation

THE INSTITUTE FOR
RESEARCH & EVALUATION

6068 S. Jordan Canal Rd.
Salt Lake City, UT 84118
ph. 801.966.5644

Re-Examining the Evidence for Comprehensive Sex Education in Schools:

A Global Research Review

February 2019

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

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.

EXECUTIVE SUMMARY

PURPOSE

To evaluate the global research evidence for school-based comprehensive sex education (CSE) according to meaningful standards of effectiveness rather than the lenient definition used by many CSE research reviews (e.g., the occurrence of one minimal positive outcome), in order to identify evidence of real program effectiveness.

BACKGROUND

The negative consequences of teenage sexual activity continue at unacceptable rates. For example, youth aged 15–24 account for 45% of all new HIV infections globally (UNESCO, 2009), and in the U.S., one in four sexually active girls has an STD (CDC, 2016). Comprehensive sex education (CSE) is widely promoted as being effective at protecting adolescents from these harms and therefore a remedy that should be implemented in school classrooms worldwide (UNESCO, 2009, 2018). Yet the permissive and explicit content of many CSE curricula raise questions about its acceptability, and the weak definitions of "effectiveness" used in many reviews of CSE research raise serious concerns about its true impact. If CSE is to be implemented on a global scale, then the question of its effectiveness in school classrooms is crucial to the real protection of youth and the prudent stewardship of public funds around the world.

METHODS

We examined the studies contained in three authoritative research reviews of sex education effectiveness: one conducted for the United Nations Educational, Scientific and Cultural Organization (UNESCO) and 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. These agencies screened several hundred sex education studies, spanning three decades, for acceptable research methods and included in their reviews only those studies that were of adequate scientific quality. There were 120 studies of school-based sex education which met that standard, including 60 U.S. studies and 43 non-U.S. studies of CSE programs (103 total) as well as 17 U.S. studies of abstinence education (AE), the often-used alternative to CSE. (The non-U.S. data did not contain enough studies of true abstinence programs for meaningful analysis.) Note: We identify a curriculum as "abstinence education" if it teaches sexual abstinence (refraining from sexual activity) as the primary protective behavior and does not promote condom or contraception use, whereas, the term "comprehensive sex education" (CSE) encompasses programs that promote condom/contraceptive use and may also teach abstinence in the same program.

We evaluated the outcomes of these 120 studies according to meaningful criteria of effectiveness grounded in the science of prevention research: effects sustained at least 12 months after the program, on a key protective indicator (abstinence, condom use—especially consistent condom use, pregnancy, or STDs), for the main (intended) teen population, based on the preponderance of research evidence and excluding programs that had any negative effects.

KEY FINDINGS

For 103 Studies of School-Based CSE Worldwide (U.S. and non-U.S combined)

OVERALL: Out of 103 school-based CSE studies worldwide (60 in the U.S., 43 internationally), only six found evidence of effectiveness (improvement on a key protective outcome—abstinence, condom use, pregnancy, or STDs—12 months after the program, for the intended population, without other negative effects). Only one of the six studies was by an independent evaluator (i.e., not the program's developer) and the results have not been replicated.

FAILURE RATE: Worldwide, school-based CSE programs that attempted to show effectiveness—sustained effects on a key protective outcome for the intended youth population—failed 87% of the time.

NEGATIVE EFFECTS: Sixteen studies (16%) found 22 instances of harmful effects by school-based CSE: increased sexual risk behavior, STDs, or pregnancy.

U.S. vs. NON-U.S.: School-based CSE programs implemented outside the U.S. appeared more likely to produce negative impact than U.S. programs: 21% of non-U.S. school-based CSE studies found harmful effects compared to 12% of the studies in the U.S. The rate of harm was 24% for school-based CSE in Africa.

PREGNANCY OR STDs: Although one of the 103 studies found a reduction in teen pregnancy and one study found a reduction in STDs, 12 months after the program for the intended population without producing other neg-

ative effects, these results have not been replicated. (Most studies did not measure these outcomes even though they are considered to be primary targets of CSE).

CONDOM USE: There was no effectiveness at increasing consistent condom use—the behavior required for meaningful protection from STDs.

DUAL BENEFIT: There was no evidence of success for the purported dual benefit of CSE: increasing both abstinence and condom use (by sexually active teens) within the same population.

For 17 Studies of School-Based Abstinence Education in the U.S.

OVERALL: Out of 17 studies of AE in the U.S., seven studies found evidence of effectiveness, an increase in teen abstinence at least 12 months after the program for the intended population, without other negative effects. Five of the seven studies were by independent evaluators, and the results have not yet been replicated.

FAILURE RATE: Of the AE programs that measured effectiveness, as defined above, 53% failed to show it.

NEGATIVE EFFECTS: One AE program (6%) produced a negative effect: an increase in number of sex partners.

PREGNANCY OR STDs: Most AE studies did not measure program effects on pregnancy or STDs and none were found. However, the increases in teen abstinence produced by seven AE programs would be expected to cause reductions in teen pregnancy and STDs, though these effects were not measured.

CONDOM USE: AE does not teach condom use and the nine studies that measured AE impact on condom use found no detrimental effects, strong evidence that AE does not do harm by reducing teen condom use.

School-based CSE Compared to AE in the U.S.

OVERALL: Seven AE studies found effectiveness compared to three studies of school-based CSE. Five of the AE studies were by independent evaluators versus none of the CSE studies. None of these results have been replicated.

FAILURE RATE: The rate of failure for school-based CSE (85%) appeared substantially higher than the rate for AE (53%).

NEGATIVE EFFECTS: For school-based sex education in the U.S., the rate of negative impact for AE appeared somewhat lower than the rate for CSE (6% vs. 12%).

SUCCESS vs. HARM: The evidence of negative effects (seven studies) appeared greater than the evidence of effectiveness (three studies) for school-based CSE in the U.S. For school-based AE in the U.S., there appeared to be more evidence of effectiveness (seven studies) than harm (one study).

CONCLUSIONS

Applying meaningful standards of effectiveness—criteria that have scientific validity and practical utility for policymakers and parents—to sex education outcomes produces a very different pattern of evidence for school-based CSE than what is typically reported by other research reviews that employ lax definitions of effectiveness. Using this more-credible approach, the claims that school-based CSE has been proven effective and AE is ineffective are not supported by 120 of the strongest and most recent outcome studies of sex education worldwide, the same studies that have been relied upon by the U.S. government and UNESCO in their extensive reviews of CSE research. *Three decades of research indicate that comprehensive sex education has not been an effective public health strategy in schools around the world, has shown far more evidence of failure than success, and has produced a concerning number of harmful impacts. The evidence for abstinence education effectiveness in the U.S., though limited, appears more promising—enough to justify additional research.*

RECOMMENDATIONS

Given the threat posed by STDs, HIV, and pregnancy to the health of young people worldwide, and the compelling lack of evidence of effectiveness for school-based Comprehensive Sex Education after nearly 30 years and 103 credible studies, we recommend that policymakers abandon plans for its global dissemination and pursue alternative prevention paradigms for reducing the negative consequences of adolescent sexual activity. Replication studies of the positive findings for abstinence education should be done to inform the development of such a paradigm.

TECHNICAL REPORT

I. Introduction

The short- and long-term consequences of teenage sexual activity continue to cause significant health and social problems in cultures and countries around the world, in spite of more than 30 years of prevention efforts. Worldwide, the AIDS epidemic continues, with "young people aged 15–24 account[ing] for 45% of all new HIV infections."[1] In the U.S., "1 in 4 sexually active adolescent females has an STD," and STD rates for adolescents are rising.[2] In addition, sexual activity for adolescents contributes to decreased mental/emotional health (e.g., higher risk of depression and suicide) and increased likelihood of sexual violence, especially for females and younger teens.[3] Moreover, the children born to unmarried teenagers are significantly more susceptible to dropping out of high school, living in poverty, and becoming teen parents themselves, in a self-perpetuating cycle.[4]

Given these harms, many public policymakers continue to place a high priority on 1) reducing teen pregnancies, 2) reducing STD and HIV infections contracted by youth, and 3) influencing adolescents to abstain from sexual activity. The wholesale delivery of "clear, well informed, and scientifically-grounded sexuality education" to youth populations worldwide is seen by many as an essential mechanism for achieving these goals in order to address the social problems at their source.[5] One type of sex education strategy promoted widely as a remedy is generally known as "comprehensive sex/sexuality education," or CSE.[6] CSE programs typically attempt to teach youth to use condoms and other contraception if they are sexually active, and if they are not, that they can choose to delay the onset of sexual activity until some indeterminate time when they are older or they decide that they are "ready."[7]

A sex education strategy often mentioned as an alternative to CSE is "abstinence education" (AE), also referred to by some as "abstinence-only" programs or "sexual risk avoidance." 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), in order to eliminate risk (rather than merely reduce it) and avoid the negative consequences of teen sex. Condom use is sometimes addressed in AE, but often in terms of its limitations or failure rates; AE does not promote or demonstrate condom or contraceptive use.[8]

The justifying rationale for CSE, and its supposed advantage over AE, has been that it is best suited to protect the full spectrum of youth from unwanted pregnancy and STDs through its purported dual benefit: that it can simultaneously increase rates of both teen abstinence (i.e., delay sexual initiation by the sexually inexperienced and promote a return to abstinence by the sexually experienced) and condom use (by teens who reject abstinence), all within the same population of youth and by a single CSE program.

However, CSE programs are often founded on a "values-free" sexual philosophy containing permissive and explicit content that can shock parents when it is revealed and is considered morally unacceptable to some, especially in more-traditional cultures.[9] Yet, because such programs are presumed to be effective, they are often presented as a necessary solution—or the only solution—to the damaging consequences of teenage sex. For example, a prominent youth advocacy organization states that CSE "has been proven effective" and that "young people need comprehensive sex education."[10] Such assumptions of CSE effectiveness are supported, if not engendered, by reports from some authoritative agencies that assert there is good scientific evidence for CSE. These are typified by statements found in the sex education "guidance" document produced by the *United Nations Educational, Scientific and Cultural Organization* (UNESCO), which asserts that abstinence programs "have been found to be ineffective" and "Programmes that combine a focus on delaying sexual activity with content about condom or contraceptive use [i.e., CSE] are effective."[11] UNESCO's "International Technical Guidance on Sexuality Education" goes on to say that "Overall, the evidence base for the effectiveness of school-based [CSE] continues to grow and strengthen, with many reviews reporting positive results on a range of outcomes" and recommends implementation of CSE programs in school classrooms worldwide as "part of the formal school curriculum," that is, to "bring CSE to children and young people everywhere."[12]

Given this focus on the school setting as a key venue for the worldwide delivery of CSE, the question of CSE effectiveness in school classrooms is crucial to the real protection of children and youth and the prudent stewardship of public funds on a global scale. Certainly the effectiveness of CSE programs should be clearly established before they are adopted and tax dollars are expended to implement them worldwide. However, weak definitions of "effectiveness" employed by many of these authoritative research reviews to evaluate CSE program outcomes raise serious questions about the real extent of CSE success.

Such concerns and the gravity of their consequences for the health of young people and for sound public policy was the impetus for our institute's examination of the best available sex education outcome research, as identified by three reputed scientific agencies, with the purpose of addressing the critical question: how effective are CSE programs in schools—what does the scientific evidence show?

II. Methods

A. Defining Program Effectiveness

We have examined many of the major reviews of sex education research conducted by key organizations in this field and have observed an important but little-reported characteristic common to many of them.[13] While most of these organizations set a reasonable standard for the quality of the *scientific methods* employed by the studies included in their review, they often employ much more lenient standards for the quality of program *outcomes* used to define effectiveness. Their claims of CSE program effectiveness are typically based on a fairly low benchmark for these outcomes, often the finding of only one minimal indicator of positive impact. This could be a short-term effect (e.g., found at three or six months but not 12 months after the program) or a subgroup effect (e.g., impact for girls but not boys) or impact on a less-protective behavior (e.g., reduced frequency of sex) while no effects are found for key protective behaviors (e.g., delayed sexual initiation or increased condom use). Often this minimal evidence comes from just one study by the program's developers (not an independent evaluator). And too often other evidence of program ineffectiveness or even harm is disregarded. This lax definition gives a different meaning to the term *effective* than what many people think of when they hear that a CSE program has "shown evidence of effectiveness."

One example is the U.S. federal *Teen Pregnancy Prevention* (TPP) initiative established by the Department of Health and Human Services in 2009 to identify evidence-based sex education programs. It designated a program as having shown "evidence of effectiveness in reducing teen pregnancy, sexually transmitted infections, and associated sexual risk behaviors" by virtue of producing only one statistically significant positive effect, even if only of short duration or only for a subgroup of the target population or found in a single study by the program's developer, and regardless of other contradictory findings.[14] Thus, two school-based CSE programs on the TPP list of evidence-based curricula (*¡Cuídate!* and *It's Your Game: Keep It Real*) actually produced no positive effects and multiple negative effects in studies by independent evaluators.[15] Yet these programs were placed on the U.S. federal TPP register as evidence based and eligible for public funding and implementation in U.S. schools because they showed some positive effects in initial studies by the programs' developers.[16] (Note: The field of prevention research cautions that study findings by program developers—who have a vested interest in the program's effectiveness—are less credible than those conducted by independent researchers. Outcome studies by program developers tend to find higher levels of effectiveness than research on the same program conducted by independent evaluators.[17] There is also a consensus in this field that programs producing both positive and negative behavioral/biological effects do not qualify for the label "effective."[18]) Thus, when brought to light, the lenient definition of effectiveness employed by some CSE research reviews can be seen to overstate or even misrepresent the scientific evidence for CSE program effectiveness—as the term is commonly understood.

The present review took a different approach: program results were evaluated according to criteria for program effectiveness derived from the field of prevention research. Assuming that adequate standards of *methodological* rigor have been met (to give confidence in the study findings), the scientific consensus on prevention research, as reflected in the work of the Society for Prevention Research, recommends defining program effectiveness according to rigorous criteria for *program outcomes* or effects.[19] We applied these recommendations in ways relevant to sex education in school settings. Specifically:

1. We looked for positive program effects (significant at the $p<.05$ level)…

    - On at least one key protective indicator (delay of sexual initiation/debut, increased condom use—especially consistent condom use, or decreased pregnancy or STDs rates),[20]
    - Sustained at least 12 months after the end of the program (thus lasting from one school year to the next),
    - Found for the main (intended) youth population, not just a subgroup,
    - Excluding programs that also produced negative effects, and
    - Based on all credible studies of the program,

including studies by independent evaluators (i.e., those who are not the program's developers or marketers).

2. Negative/harmful program effects on important sexual health indicators were counted if they impacted the intended population or a substantial subgroup (e.g., males only or females only) and lasted for any duration. Such negative program impacts are a cause for concern and negate a prevention program's claim to "effectiveness," according to a consensus in the field of prevention program research.[21]

Applying these more-credible standards of effectiveness to CSE program outcomes enabled us to identify *meaningful* evidence of CSE program effectiveness, evidence that has scientific validity and practical utility for policymakers and parents.

A note about consistent condom use (CCU): consistent condom use (i.e., using a condom with every act of sexual intercourse) is required for effective condom protection. According to the *Centers for Disease Control and Prevention*, "Consistent and correct use of male latex condoms can reduce (though not eliminate) the risk of STD transmission. To achieve the maximum protective effect, condoms must be used both consistently and correctly. Inconsistent use can lead to STD acquisition because transmission can occur with a single act of intercourse with an infected partner."[22] This is illustrated by a study of African American teenage girls that found 17.8% of those who used condoms consistently acquired an STD, but the number was 30% for those who used condoms inconsistently.[23] At least three peer-reviewed studies have found STD rates were higher for inconsistent condom users than non-users.[24] (Even consistent condom use does not provide the 100% protection from STDs afforded by abstinence,[25] nor prevent the increased emotional harm and sexual violence associated with teen sex.[26])

However, most CSE studies do not measure CCU but instead assess less-protective indicators—frequency of condom use or use at last intercourse. This review distinguished between measures of "consistent condom use" (CCU) and "less-protective measures of condom use," and reported research findings for both. However, where both were measured in the same study, the CCU outcome was considered the key indicator, with failure on this outcome not outweighed by success on a less-protective measure of condom use. On the other hand, where CCU was *not* measured, we accepted a less-protective measure of condom use as an indicator of program effectiveness.

B. The Database

Many hundreds of studies of sex education program effectiveness have been conducted in the U.S. and worldwide since such programs became popular in the early 1990s. This large universe of studies has been reviewed and sifted by many 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 agencies: the *Teen Pregnancy Prevention program* (TPP) in the U.S. Department of Health and Human Services (HHS),[27] the *Community Preventive Services Task Force* supported by the *U.S. Centers for Disease Control & Prevention* (CDC),[28] and the *United Nations Educational, Scientific and Cultural Organization* (UNESCO).[29] Each of these entities has identified and reviewed the credible studies of CSE conducted since 1990. (For the two U.S. agencies, their reviews covered only sex education implemented in the United States, while the UNESCO review included programs in both U.S. and non-U.S. settings.) The TPP review produced a list of CSE programs that it declares "have [shown] evidence of effectiveness" while the other two agencies have stated, based on their reviews, that CSE has shown sufficient evidence of effectiveness in school settings to recommend it as a prevention strategy.[30]

Because the studies included in these three databases met the standards for adequate research quality established by these preeminent agencies, and because our focus was programs in school settings, we used the studies of *school-based* sex education contained in these three reviews as the database for our analysis. This allowed us to examine what other experts have independently identified as some of the best evidence for school-based CSE effectiveness. (Note: We defined a sex education program as "school- based" if it occurred in a school classroom during the normal school day, or recruited its subjects from the school population and occurred after school or at the school on Saturdays, and the majority of the program was not community-based.)

Combining these three reviews yielded 103 studies of 79 CSE programs[31] in school settings around the world: 60 studies of 40 programs in the U.S. and 43 international studies of 39 programs in other countries (40 of the non-U.S. studies were in "low- or middle-income" countries, including 29 in Africa). In addition, there were 17 studies of 16 school-based abstinence education programs (AE) conducted in the U.S. that had met the same standards of research quality and were included in the same database. (Note: The international data did not contain enough studies of true abstinence-only programs for meaningful analysis.) This provided a total of 120 studies for our review.[32]

We examined each of these studies, rather than relying on summaries by other reviewers, and evaluated the programs' outcomes according to the criteria outlined above. Our results are summarized in Tables 1 – 4 below and shown study by study in Tables 5 – 7.

III. Results

Using criteria for program effectiveness derived from the field of prevention science—criteria that are more rigorous than the lenient standards of effectiveness often employed in other reviews of CSE outcome research—produced different findings than what has typically been reported by such reviews, findings that do not support the claim that CSE in school settings has been proven effective, and AE has been proven ineffective.

A. Findings for U.S. School-Based Comprehensive Sex Education

For the 60 studies of 40 school-based CSE programs in the U.S., three studies, representing three programs, found positive impact at least 12 months after the program on a key protective outcome for the intended population without other negative effects. None of the three studies was conducted by an independent evaluator (i.e., someone other than the program developer or marketer), and replication studies have not confirmed the initial positive results. In contrast to the positive effects, seven studies of six programs found harmful CSE program impact: increased sexual risk behavior or reduced sexual health.

PREGNANCY or STDs. None of the 40 school-based CSE programs showed reductions in teen pregnancy beyond the end of the program, and none reduced STDs. (Few programs even measured these outcomes.) One CSE program actually increased teen pregnancy for females in a school-based population.[33]

ABSTINENCE/SEXUAL INITIATION. Only one school-based CSE program showed effectiveness at increasing teen abstinence (i.e., delaying sexual initiation). However, the study was by the program developer, and evidence from multiple replication studies did not confirm the original positive results.[34]

CONSISTENT CONDOM USE. There was no evidence of school-based CSE effectiveness at producing sustained increases in *consistent condom use* by teens. (Consistent use is necessary to provide optimum protection from STDs.) One school-based CSE program reported a sustained effect in a study by its developer, but a study by an independent evaluator did not confirm that effect and found that the CSE program increased sexual risk behaviors for major subgroups of the target population.[35]

FREQUENT OR RECENT CONDOM USE. Among programs that did not measure consistent condom use, two showed sustained increases in less-protective measures of condom use (e.g., frequent or recent use) for the intended population. But the studies were by program developers, and the findings have not been replicated.[36]

DUAL BENEFIT. There were no increases in both teen abstinence and condom use (by sexually active teens) within the same CSE program and teen population at least twelve months after the program's end.

PROGRAM SUCCESS VERSUS FAILURE. Only 20 of the 40 programs actually measured the more-rigorous definition of effectiveness (i.e., protective impact on a key indicator, at least 12 months post-program, for the intended population, without other negative effects), and only three met that standard. (None of the three studies were by independent evaluators, and two of them used the weaker outcome measure of "more frequent" or "at last intercourse" condom use.) This was a success ratio of 15% (3/20), or, inversely, 85% of the CSE programs in U.S. schools that measured real effectiveness failed to demonstrate it.[37]

HARMFUL PROGRAM IMPACT. Seven studies reported ten findings of harmful impact, on the main (intended) population or a substantial subgroup, produced by six school-based CSE programs (some programs produced multiple negative effects): three increased rates of recent sex, one increased sexual initiation, two increased oral sex, one increased teen pregnancy, one increased number of sex partners, and two reduced condom/contraceptive use.[38] This was 12% (7/60) of the studies or 15% of the 40 school-based CSE programs (6/40) that showed harmful effects, which are higher rates than would be expected by chance (5%).

B. Findings for U.S. School-Based Abstinence Education

The 17 studies of 16 school-based abstinence education programs in the U.S. found that seven AE programs

delayed sexual initiation (increased abstinence) at least 12 months after the program for the intended population, without other negative effects, and five of these seven studies were by independent evaluators. These results have not yet been replicated. The nine studies that measured condom use found no detrimental effects. Only one AE program showed a negative program effect: an increase in number of sex partners.

ABSTINENCE/SEXUAL INITIATION. Seven school-based abstinence education (AE) programs produced sustained (12-month post-program) delays in teen sexual initiation (increased rates of abstinence) for the intended population.[39] Five of the seven studies were by independent evaluators.[40] Three of the seven programs also produced a sustained reduction in frequent or recent sex, a move toward abstinence by sexually experienced teens.[41] Only two replication studies have been conducted of these results: a second study of one program only measured short-term effects and found some;[42] a second study of another program found inconclusive results.[43]

CONSISTENT CONDOM USE. AE does not promote condom use so it would not be expected to produce improvement on this outcome. Five studies measured consistent condom use and found no significant effect.

ANY CONDOM USE. A total of nine studies tested AE impact on condom use (whether consistent, frequent, or recent use) with none finding a negative effect[44] and one AE program producing an increase in condom use frequency12 months after the program.[45]

TEEN PREGNANCY OR STDs. There was not adequate evidence about AE impact on pregnancy or STDs. Only four AE programs in the database measured these outcomes, but none of the four found impact on abstinence, so it was not surprising that there were also no effects on pregnancy or STDs. (In fact, the evaluation studies of these four programs had some methodological/design problems that raise questions about their results.[46]) However, the increases in teen abstinence caused by seven other AE programs would be expected to produce reductions in teen pregnancy and STDs, though unmeasured.

DUAL BENEFIT. AE would not be expected to improve condom use and none of the programs produced sustained increases in both abstinence and condom use (by the sexually active).

PROGRAM SUCCESS VERSUS FAILURE. Of the 15 AE programs that measured effectiveness, as defined previously, seven met that standard, for a success ratio of 47% (7/15). Inversely, 53% of AE programs in U.S. school settings that measured effectiveness failed to produce it.

HARMFUL PROGRAM IMPACT. One of the 17 AE studies (6% of the programs/studies) reported a negative effect: an increase in number of sex partners.[47]

C. Findings for International School-Based Comprehensive Sex Education

Of the 43 studies that evaluated 39 school-based CSE programs outside the United States, three programs produced positive impact 12 months after the program, on a main protective outcome, for the intended population, without other negative effects. Only one of the three studies was by an independent program evaluator, and none of the results have been replicated. Nine international studies found harmful CSE effects.

PREGNANCY OR STDs. Only one of the 39 school-based CSE programs in a non-U.S. country showed effectiveness (as defined above) at reducing teen pregnancy, in a study by independent evaluators.[48] Only one study (by the program's developer) found effectiveness at reducing STDs.[49] Very few studies measured (or reported) program effects on teen pregnancy or STDs, even though reducing these harms is a central purpose of the CSE strategy.

ABSTINENCE/SEXUAL INITIATION. Only one of the school-based CSE programs in a non-U.S. setting showed effectiveness at delaying teen sexual initiation.[50] The study was by the program developer, and the effects have not been replicated.

CONSISTENT CONDOM USE. None of the school-based CSE programs in non-U.S. countries showed an increase in consistent condom use for any period of time or any subgroup; very few studies (9) even measured this outcome. (Consistent condom use is necessary for optimum protection from STDs.)

FREQUENT OR RECENT CONDOM USE. Only one of the school-based CSE programs in a non-U.S. setting showed an increase in a less-protective measure of condom use (recent use) 12 months after the program for the intended population and without negative effects on other outcomes. But because the same study also measured consistent condom use—the more-protective outcome—without finding significant impact, the effect on the less-protective measure was not counted here as evidence of program effectiveness.[51]

DUAL BENEFIT. None of the school-based CSE programs in a non-U.S. setting showed effectiveness at achieving the dual benefit intended by most CSE programs—a sustained increase in both teen abstinence and condom use (by the sexually active) for the intended population within the same CSE program.

PROGRAM SUCCESS VERSUS FAILURE. Out of the 27 non-U.S. programs that actually measured effectiveness (impact on a key outcome, at least 12 months post-program, for the intended population, without other negative effects), only three met that standard (one program reduced teen pregnancy, one reduced STDs, and one delayed sexual initiation), a success ratio of 11% (3/27). Inversely, 89% of international school-based CSE programs that measured evidence of effectiveness failed to demonstrate it.

HARMFUL PROGRAM IMPACT. Nine school-based CSE programs in non-U.S. settings caused 12 negative impacts (i.e., did harm to program participants): they either increased teen sexual initiation, STDs, number of partners, recent sex, paid sex, or forced/coerced intercourse, or they decreased condom use.[52] Three of these programs had harmful impacts on multiple outcomes.[53] Thus, one in five school-based CSE programs outside the U.S. produced negative effects (9/39 programs, 23%, or 9/43 studies, 21%).

EFFECTS IN AFRICA. Within the database was a subset of 29 studies of school-based CSE in Africa, representing 26 different programs. Of these, 19 measured CSE program impact after 12 months, with two showing effectiveness on one of the key protective indicators (one reduced STDs[54] and one delayed sexual initiation [55]), for a success ratio of 11% (2/19). Inversely, 89% of African school-based CSE programs that measured evidence of effectiveness failed to demonstrate it. Seven of the 29 African studies (24%), examining 26 programs, found negative impacts.[56] Thus, 27% (7/26) of the African school-based CSE programs produced negative effects.

D. Global Findings for School-Based CSE (U.S. and non-U.S combined)

Of the 79 U.S. and international school-based CSE programs evaluated by 103 studies, six studies of six programs found sustained improvement on one of the key protective outcomes, for the intended population, without other negative effects.[57] Only one of the studies was by an independent evaluator.[58] There was no effectiveness at increasing consistent condom use or at achieving the purported dual benefit of CSE: increased abstinence and condom use within the same program. In addition, sixteen studies of 15 programs found negative CSE effects: increased sexual risk behavior, STDs, or pregnancy.

PREGNANCY OR STDs. Worldwide, one out of 79 school-based CSE programs reduced teen pregnancy and one reduced STDs, 12 months after the program, for the intended population, without producing other negative effects.

ABSTINENCE/SEXUAL INITIATION. Worldwide, two out of 79 school-based CSE programs reduced teen sexual initiation, 12 months after the program, for the intended population, without producing other negative effects.

CONSISTENT CONDOM USE. Worldwide, none of the 79 school-based CSE programs produced an increase in consistent condom use by adolescents, 12 months after the program, for the intended population, without producing other negative effects. (Consistent condom use is necessary for optimum STD protection.)

FREQUENT OR RECENT CONDOM USE. In the absence of a measure of consistent condom use, two of the 79 school-based CSE programs worldwide produced an increase in less-protective measures of condom use, 12 months after the program, for the intended population, without producing other negative effects.

DUAL BENEFIT. Worldwide, none of the 79 school-based CSE programs showed effectiveness at achieving the dual benefit intended by most CSE programs—a sustained increase in both teen abstinence and condom use (by the sexually active) for the intended population within the same CSE program.

PROGRAM SUCCESS VERSUS FAILURE. A global "success ratio," estimated by taking the six programs that *produced effectiveness* as a proportion of the 47 school-based CSE programs worldwide that *measured effectiveness* (i.e., a 12-month post-program effect on one of the key indicators), was six out of 47 or 13%. Inversely, 87% of the school-based CSE programs worldwide that measured effectiveness failed to show it.

HARMFUL PROGRAM IMPACT. Of the 103 school-based CSE studies worldwide, a total of 16 studies that evaluated 15 programs found 22 instances of negative/harmful CSE impact on teen sexual health or risk behavior (six programs produced multiple negative effects).[59] This was 16% (16/103) of the studies or 19%

(15/79) of the school-based CSE programs globally that showed negative impact, proportions which are both higher than would be expected by chance.

E. Findings by Region (U.S. vs. Non-U.S.) and Program Type (CSE vs. AE)

These findings are shown by outcome, geographic region, and program type in Tables 1-4.

INDEPENDENT EVIDENCE. As shown in Table 1, the evidence of effectiveness for school-based CSE that came from independent studies—those not conducted by the program's developers or marketers—was very small and only found in a non-U.S. setting: one study found a positive impact at least 12 months after the program for the intended population on key protective outcomes, without producing other negative effects, in a study by independent evaluators. This compares to five studies of school-based AE in the U.S. that met this standard.

PROGRAM SUCCESS VERSUS FAILURE. As shown in Table 2, the success ratio appeared somewhat similar for school-based CSE in U.S. settings (15%) and outside the U.S. (11%). By comparison, the smaller number of studies of U.S. school-based abstinence education (AE) showed a substantially higher success ratio of 47%. The inverse of these numbers, indicating a rate of program failure, were 85% for school-based CSE in the U.S., 89% for these programs in non-U.S. settings, and 53% for AE programs in the U.S.

HARMFUL PROGRAM IMPACT. School-based CSE programs implemented outside the U.S. appeared more likely to produce negative impact than U.S. programs (see Table 3). In the U.S., 12% of studies (7/60) found negative effects by six programs (6/40 or 15% of school-based CSE programs), while outside the U.S., 21% of studies (9/43) found negative effects for school-based CSE (9/39 or 23% of programs). The majority of non-U.S. studies took place in Africa (29 out of the 43 studies), where the rate of negative impact appeared even higher (24% of studies, 27% of programs). For the 17 studies of school-based AE in the U.S., negative impact was found for one program, which was about 6% of the programs/studies.

EVIDENCE OF EFFECTIVENESS VERSUS HARM. Another way to summarize these findings is to compare the amount of evidence of program effectiveness/success to the amount of evidence of negative/harmful impact. Worldwide, in terms of sheer quantity, there appeared to be more evidence of harm by school-based CSE, 16 studies, than evidence of real effectiveness, six studies. The pattern was seen for school-based CSE both within and outside the U.S., but was reversed for school-based AE in the U.S., with more evidence of effectiveness, seven studies, than harm, one study (see Table 4).

Another way to assess this difference would be to look at comparative rates of impact, that is, percentages of programs showing effectiveness/success versus percentages producing harmful impact. However, this was not done because it did not appear to be scientifically defensible, that is, a type of apples to oranges comparison.[60]

IV. Limitations

This review was not a statistical meta-analysis in which study outcomes are combined numerically and statistically significant differences can be calculated. (The meta-analysis methodology is most appropriate when program methods, settings, and populations are homogenous, and this was not the case with our sample of CSE programs. There was a high degree of heterogeneity on those categories across programs, even though they were all school based, such that a meta-analysis would not have been useful.) Therefore, where comparisons were made across types of outcomes, by geographic region, or by program type, they were estimates meant to identify a pattern of evidence; we did not conduct statistical tests of differences. This study also did not report on the size of program effects in terms of the amount of behavioral change or the percentage of participants impacted, but rather looked to the statistical significance and duration of effects to identify important program outcomes. Finally, within this database of 120 studies—each of which had been vetted for adequate research rigor by at least one of three credible scientific agencies (UNESCO, CDC, HHS)—there was still meaningful variation between studies in the quality of the scientific methods employed. For example, even among randomized controlled trials (the strongest type of study) we saw the use of weak statistical analyses, study design problems that could undermine the detection of effects, as well as sizable pre-test differences between groups that were not controlled for in the post-test results. This observation underscores the need for stronger evidence about sex education effectiveness in school classrooms.

V. Discussion

This review demonstrates the value of employing credible criteria—standards that provide a useful real-world definition of program effectiveness, grounded in the scientific field of prevention research—when evaluating sex education success. Applying such criteria to school-based programs worldwide, both within the United States and internationally, we found very little evidence of CSE effectiveness in school settings—there was far more evidence of CSE failure (87%) than success (13%).

Our analysis paints a very different picture than the reports of success presented by other reviews of CSE research. Some of these have looked at the same studies but used a more-lenient, less-credible definition of effectiveness when evaluating program outcomes. Some have also mixed school-based results in with those of clinic- and community-based programs, where the methods differ and the program outcomes are somewhat better. In light of UNESCO's goal to implement CSE in schools globally, we expect that the discrepancy between our finding of little school-based CSE effectiveness and the CSE success typically reported by other reviews will be of interest to policymakers concerned with protecting children.

Ironically, the evidence cited by three reputable agencies—UNESCO, CDC, and HHS—to support their assertions that school-based CSE programs are effective appears to undermine those claims:

- UNESCO states that "Overall, the evidence base for the effectiveness of school-based [CSE] continues to grow and strengthen, with many reviews reporting positive results on a range of outcomes."[61]

- The CDC-supported meta-analysis asserted that CSE programs are effective "across a range of populations and settings … [including] both … school and community settings."[62]

- The U.S. Health and Human Services *Teen Pregnancy Prevention* website indicates that all of the school-based CSE programs on its list have "shown evidence of effectiveness."[63]

Yet the findings from the 103 school-based CSE studies in their combined databases do not support these assertions. Out of the 103 school-based CSE studies, only six found evidence of real effectiveness: protective impact at least 12 months after the program for the intended population without producing other negative effects. Notably, there was no evidence of success at increasing consistent condom use—the behavior required for significant protection from STDs—and no evidence of success at the dual benefit that is the supposed hallmark of the CSE approach: increasing both teen abstinence and condom use within the same population.

The fact that almost all of the evidence of school-based CSE effectiveness (5 out of 6 studies) was produced by the programs' developers should not be taken lightly. For example, approximately one-half of the 60 U.S. school-based CSE studies were by program developers,[64] and these studies were about twice as likely as the studies by independent evaluators (57% compared to 28%) to report any positive program outcomes, that is, when not defined by the higher standards of effectiveness employed in the present study. When using these higher standards to count only evidence of real effectiveness (12-month post-program effects on the intended population, etc.) all of the U.S. studies by independent evaluators dropped out, leaving three studies by program developers. Some have argued that while this pattern could be due to bias by program developers, it could also be influenced by superior implementation of programs by their developers, which would produce better study outcomes.[65] However, we did see evidence of possible researcher bias in some of the studies by program developers that we reviewed.

Perhaps of greatest concern, the six studies that did find some evidence of school-based CSE effectiveness stand in contrast to the 16 studies that found 22 negative effects on teen sexual health and risk behavior. There were 18 increases in teen sexual activity or other risk behaviors, in direct contradiction to UNESCO's assertion that CSE "does not increase sexual activity [or] sexual risk-taking behaviour."[66] In fact, there was a concerning number of harmful effects on program participants (22), and a concerning prevalence of harmful impact: 16% of studies (16/103, nearly one in six) or 19% of school-based CSE programs (15/79, nearly one in five). In terms of quantity of evidence (i.e., number of studies), CSE programs in school classrooms worldwide appear to have produced more evidence of harm (16 studies) than of real effectiveness (six studies). The rate of negative impact was especially high for CSE programs in African schools, where it was approximately one in four studies/programs, a finding that is even more serious in light of the fact that Africa continues to be the continent most impacted by HIV and AIDS.

Finally, the scientific evidence reported here contradicts the oft-repeated claim that research shows abstinence education (AE) is ineffective and/or harmful. (See for example, this statement by UNESCO, "Programmes that promote abstinence-only have been found to be ineffective

in delaying sexual initiation, reducing the frequency of sex or reducing the number of sexual partners … and [are] potentially harmful to young people's sexual and reproductive health and rights."[67]). Seven studies in this database—studies found to be of adequate scientific rigor by either UNESCO, the CDC, or HHS—demonstrated a long-term delay in sexual initiation, and three of these also produced long-term reductions in sexual activity by sexually experienced teens (see Table 6). The rate of AE effectiveness/success was nearly one out of two (47%), and the prevalence of harmful effects, at 6%, was about what would be expected by chance. Furthermore, none of the nine studies that tested AE impact on condom use found negative effects. This strong evidence contradicts the charge that AE does harm by reducing condom use.

It will no doubt come as a surprise to many that this credible database produced better evidence for the effectiveness of AE than for CSE in U.S. schools. This is especially noteworthy considering the markedly fewer number of available AE studies, and the fact that, unlike the CSE results, most of the AE evidence was produced by independent evaluators (not program developers). The *amount* of AE evidence of effectiveness, in terms of number of studies, appeared somewhat greater than for CSE in U.S. schools (seven AE studies vs. three CSE studies) and the overall success *rate* for AE programs, at 47%, appeared to be much higher than that of school-based CSE in the U.S., at 15%. Moreover, the prevalence of negative effects appeared somewhat lower for AE (6%) than for CSE in U.S. schools (12%).

It is worth noting that the seven AE programs which increased teen abstinence after 12 months appear to have provided total protection for those youth during that time, by their avoidance of sexual risk behavior. Only two of the 79 school-based CSE programs in this worldwide database provided this protection by increasing teen abstinence after 12 months, without other negative effects. Nevertheless, it should also be noted that the AE database reviewed was small and limited to studies in the U.S., thus, it is not adequate to support the drawing of firm conclusions. Additional studies should be done in order to expand the AE evidence base and to determine if the positive AE findings are replicable.

We end with an observation about program *potential* versus program *effectiveness*. It is not difficult to find sex education programs that have only produced results on less-protective outcomes, or for short durations, or only for subgroups of the intended population. While such outcomes can identify programs that *may have potential*, according to the field of prevention research, this is not sufficient evidence of effectiveness to justify widespread dissemination in schools, nor financial support using public funds. Some programs in this database showed evidence of potential by producing effects that approached the cut-off points for our criteria of effectiveness. (These outcomes are highlighted in blue shading in Tables 5 – 7.) However, better results than these are needed to justify designation as a truly effective program that can be utilized with confidence. Such a conclusion is consistent with the findings of *Blueprints for Healthy Youth Development,* a reputable registry of evidence-based prevention programs covering the spectrum of youth risk behaviors. Based on its review of the research evidence, as of this printing, *Blueprints* has not named any school-based CSE program as a "Model Program" and lists only five as "Promising."[68] According to the *Blueprints* website, only Model Programs "are deemed ready for widespread use."[69]

VI. Conclusions

When measured by credible criteria derived from the field of prevention research, a database containing 103 of the strongest and most recent CSE studies, vetted for research quality by three reputed scientific agencies (UNESCO, CDC and HHS), shows very little evidence of CSE effectiveness in school populations and settings. Where there was some evidence, nearly all of it was produced by the program's developers and had not been replicated. Thus, three decades of research indicate that CSE has not been an effective public health strategy in classrooms around the world and that too many programs may be doing harm. When applying the same standards of effectiveness to AE in U.S. schools, the evidence—though limited—is more independent and looks more promising than the results for CSE, enough to justify funding additional AE research.

VII. Recommendations

Given the threat posed by STDs, HIV, and pregnancy to the health and well-being of young people worldwide, and the compelling evidence of ineffectiveness by school-based Comprehensive Sex Education after three decades of research, policymakers should abandon plans for its global dissemination and pursue alternative prevention paradigms to prevent the negative consequences of adolescent sexual activity. Replication studies on the promising results for Abstinence Education in the U.S. should be done to inform the development of such a paradigm.

ENDNOTES

1. 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

2. 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

3. Hallfors, D. D., Waller, M. W., & Ford, C. A., Halpern, C. T., Brodish, P. H., Iritani, B. (2004). Adolescent depression and suicide risk: association with sex and drug behaviors. *American Journal of Preventive Medicine 27*(3)*,* 224–231. doi:10.1016/j.amepre.2004.06.001; 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. doi.org/10.1016/j.jhealeco.2008.05.008; Meier, A. M. (2007). Adolescent first sex and subsequent mental health. *American Journal of Sociology, 112*(6), 1811-1847. doi.org/10.1086/512708; Else-Quest, N. M., Hyde, J. S., DeLamater, J. D. (2005). Context counts: Long-term sequelae of premarital intercourse or abstinence. *Journal of Sex Research, 42,* 102e12. doi.org/10.1080/00224490509552263; Kramer, A. (2014). Virgin Territory: What Young Adults Say About Sex, Love, Relationships, and the First Time. *The National Campaign to Prevent Teen and Unplanned Pregnancy*. Retrieved from https://thenationalcampaign.org/sites/default/files/resource-primary-download/virgin-territory-final.pdf; 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. doi:10.1542/peds.114.2.e220.

4. Wilcox, B. W. (2009, May 22). The Real Pregnancy Crisis. *The Wall Street Journal.*; Ravitz, J. (2009, April 8). Out-of-wedlock births hit record high. Retrieved from http://www.cnn.com/2009/LIVING/wayoflife/04/08/out.of.wedlock.births/index.html; "Teen Pregnancy" retrieved from https://powertodecide.org/what-we-do/information/why-it-matters

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

6. Comprehensive sex education (CSE) is sometimes presented under different labels/names, often under the banner of comprehensive sexuality education. For example, in the Caribbean region, CSE programs are often promoted as Home and Family Life Education (HFLE). Internationally, CSE is also often promoted using the labels of sexual and reproductive health education or teen pregnancy or HIV prevention education.

7. One popular CSE program, ¡*Cuídate!*, 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?" (see Villarruel, A. M., Jemmott, L. S., Jemmott, J. B. (n.d.). *Facilitator's Curriculum, Module 2: Building Knowledge About Pregnancy, STDs and HIV*).

8. We identified a curriculum as "abstinence education" (or AE) if it teaches abstinence/sexual risk avoidance as the primary protective behavior and does not promote condom or contraception use as a risk reduction alternative. We applied the "comprehensive sex education" (CSE) label to programs that promote condom and/or other contraceptive use and may also encourage abstinence (to some degree). However, within these two categories of interventions is a spectrum of programs that vary in the relative amount of emphasis given to the respective topics.

9. For example, 1) UNESCO's *Technical Guidance on Sexuality Education*, *An Evidence-Informed Approach*, Revised Edition, 2018 (see: http://www.unaids.org/sites/default/files/media_asset/ITGSE_en.pdf) recommends content for CSE programs worldwide. The document considers children as young as nine years old to be "able to make informed decisions about sexual behaviour, including whether to delay sex or become sexually active" and directs that they be taught to "describe male and female responses to sexual stimulation" and "understand that abstinence means choosing not to have sex, or deciding when to start having sex and with whom …" (p. 71). It recommends teaching children as young as 12 years old "that non-penetrative sexual behaviours are without risk of unintended pregnancy, offer reduced risk of STIs, including HIV, and can be pleasurable" (p. 72). In a section directed at the same age group, UNESCO discusses "transactional sexual activity"—giving sex in exchange for money or favors—in values-neutral terms. Rather than teaching that paid sex is inherently harmful to them and should be unequivocally avoided, children as young as 12 are to be told that paid sex "can pose risks to one's health and well-being … [and] can increase vulnerability and limit the power to negotiate safer sex" (p. 72); 2) The CSE curriculum "Making Proud Choices," which targets children ages 11-13 years old, "teaches participants… ways to make condoms pleasurable" (see: https://tppevidencereview.aspe.hhs.gov/document.aspx?rid=3&sid=102&mid=2);); 3) in the CSE curriculum "Be Proud Be Responsible," young people ages 13 and up are taught that "sex is still fun and pleasurable when a condom is used" and "are taught how to incorporate this belief into role-play scenarios" (see: https://tppevidencereview.aspe.hhs.gov/document.aspx?rid=3&sid=21&mid=2) and the lessons include "condom demonstrations and other exercises" (see: https://www.etr.org/ebi/programs/be-proud-be-responsible/).

10. Advocates for Youth. (2009). Comprehensive Sex Education: Research and Results. *The Facts, September 2009*. Retrieved from https://www.advocatesforyouth.org/wp-content/uploads/storage//advfy/documents/fscse.pdf

11. United Nations Educational, Scientific and Cultural Organization. (2018). *International Technical Guidance on Sexuality Education: An Evidence-Informed Approach*, Revised Edition. p. 18. Retrieved from http://www.unaids.org/sites/default/files/media_asset/ITGSE_en.pdf

12. United Nations Educational, Scientific and Cultural Organization. (2009). *International Technical Guidance on Sexuality Education, Volume 1*. p. iii; United Nations Educational, Scientific and Cultural Organization. (2018). *International Technical Guidance on Sexuality Education: An Evidence-Informed Approach*, Revised Edition, p. 12, 28. Retrieved from http://www.unaids.org/sites/default/files/media_asset/ITGSE_en.pdf

13. For example, *The National Campaign to Prevent Teen and Unplanned Pregnancy* (now called *Power to Decide: the Campaign to Prevent Unplanned Pregnancy*), *Child Trends*, the *U.S. DHHS Teen Pregnancy Prevention Program*, the *Centers for Disease Control and Prevention*, and *UNESCO*.

14. Teen Pregnancy Prevention Evidence Review. (n.d.). Retrieved from https://tppevidencereview.aspe.hhs.gov/EvidencePrograms.aspx

15. Kelsey, M., Layzer, C., Layzer, J., Price, C., Juras, R., Blocklin, M., Mendez, J. (2016). Replicating ¡Cuídate!: 6-Month Impact Findings of a Randomized Controlled Trial. *American Journal of Public Health, 106*(S1), S70–S77. doi:10.2105/AJPH.2016.303371; 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., doi.org/10.1016/j.jadohealth.2013.10.204; 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. doi:10.2105/AJPH.2016.303419; 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. Retrieved from https://www.hhs.gov/ash/oah/sites/default/files/ash/oah/oah-initiatives/evaluation/grantee-led-evaluation/reports/uthsc-final-report.pdf

16. 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. doi:10.1001/archpedi.160.8.772; Tortolero, S., Markham, C., Fleslcher, M., Shegog, R., Addy, R., Escobar-Chaves, L., Baumler, E. R. (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. doi.org/10.1016/j.jadohealth.2009.06.008; 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. doi.org/10.1016/j.jadohealth.2011.07.010

17. For example, the *Society for Prevention Research* 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 "Standards for Effectiveness: 5f" 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.org/10.1007/s11121-015-0555-x. Retrieved from http://www.preventionresearch.org/wp-content/uploads/2011/12/Standards-of-Evidence_2015.pdf). This concern was also 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" (See: Farb, A. & Margolis, A. (2016). The Teen Pregnancy Prevention Program (2010–2015): Synthesis of Impact Findings. *American Journal of Public Health, 106* (Suppl 1). doi: 10.2105/AJPH.2016.303367.).

18. 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 "a consistent pattern" of positive effects and "no serious negative (iatrogenic) effects on important outcomes" both within the same study and across multiple evaluation studies (see: "Standards of Efficacy: 7b and 7c" in 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. Retrieved from http://www.preventionresearch.org/wp-content/uploads/2011/12/Standards-of-Evidence_2015.pdf, p. 910); the *Blueprints for Healthy Youth Development* website stipulates that to be considered even a "Promising Program" (the less-stringent label a step below their "Model Program" designation) there should be "no evidence of harmful [program] effects" (See: https://www.blueprintsprograms.org/criteria).

19. The development of standards for scientific evidence of program effectiveness has been undertaken by national entities like *The Society for Prevention Research (SPR)* and *Blueprints for Healthy Youth Development* (https://www.blueprintsprograms.org/criteria). 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. doi.org/10.1007/s11121-005-5553-y), 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.org/10.1007/s11121-015-0555-x. These standards include a requirement of long-term post-program effects as well as a concern about main effects vs. subgroup effects, the importance of evidence from multiple studies, including those by independent evaluators (not the program developers), and the stipulation that significant negative program effects disqualifies an intervention from designation as "effective." For a more complete discussion/documentation of these standards, see the full report of our review of U.S. studies: Weed, S., & Ericksen, I. (2018). Re-examining the Evidence for Comprehensive Sex Education in Schools: Part One – Research Findings in the United States. Salt Lake City: *The Institute for Research & Evaluation*. Retrieved from http://www.institute-research.com/CSEReport/Reexamining_the_Evidence-CSE_in_USA_5-29-18FINAL.pdf

20. In the body of evidence reviewed here, about one-third of the school-based CSE studies did not measure condom use, while some measured "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 (see Ralph, L. J., McCoy, S. I., Shiu, K., Padian, N. S. (2015). Hormonal contraceptive use and women's risk of HIV acquisition: a meta-analysis of observational studies. *The Lancet Infectious Diseases, 15,* 181–189. doi.org/10.1016/S1473-3099(14)71052-7; Li, C. I., Beaber, E. F., Tang, M. C., Porter, P. L., Daling, J. R., Malone, K. E. (2012). Effect of Depo-Medroxyprogesterone Acetate on Breast Cancer Risk among Women 20 to 44 Years of Age. *Cancer Research, 72,* 2028–2035. doi**:**10.1158/0008-5472.CAN-11-4064). Another commonly used program outcome measure is to ask teens if they have had "unprotected sex," where a "no" response means they have either used any one of many 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 obscures what the program's protective effect really was, especially whether it protected teens from STDs and HIV through increased abstinence or consistent 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 indicators of program effectiveness. In spite of this inadequacy, we reported on "unprotected sex" when it appeared to be a type of surrogate measure for condom use. With regard to measures of sexual activity/abstinence, the primary goal for programs targeting youth populations should be to delay the initiation or onset of sexual activity; thus, this measure was considered a key indicator of program effectiveness. Some studies also measured recent sexual activity (usually for the past three months), indicating a possible return to abstinence. However, if there was no program impact on sexual initiation, then an effect on recent sex—especially for such a brief time span—while certainly desirable, was not considered sufficient evidence of program effectiveness. The same argument was applied to measures of frequency of sex or number of sex partners, both of which—while they may *reduce* risk—can still leave youth at substantial risk for STDs and pregnancy.

21. See Endnote #17.

22. Centers for Disease Control and Prevention. *Condoms and STDs: Fact Sheet for Public Health Personnel.* Retrieved from https://www.cdc.gov/condomeffectiveness/latex.html

23. 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. doi: 10.2105/AJPH.93.6.901.

24. 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" as compared to nonuse (Ahmed, S., Lutalo, T., Wawer, M., Serwadda, D., Sewankambo, N. K., Nalugoda, F., Makumbi, F., Wabwire-Mangen, F., Kiwanuka, N., Kigozi, G., Kiddugavu, M., Gray, R. (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., Douglas, J. M. (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. doi:

10.1097/01.OLQ.0000114338.60980.12). And a study in Brazil found that condom use at last intercourse (this is not a measure of consistent use) was independently and positively related to infection with a high-risk HPV type as compared to nonuse (OR 1.3, p<.05) (see: Grinsztejn, B., Veloso, V., Levi, J., Velasque, L., Luz, P., Friedman, R. K., Andrade, A. C., Moreira, R. I., Russomano, F., Pilotto, J. H., Bastos, F. I., Palefsky, J. (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. doi.org/10.1016/j.ijid.2008.03.031).

25. 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., Stanberry, L., Kamb, M., Wald, A. (2009). A Pooled Analysis of the Effect of Condoms in Preventing HSV-2 Acquisition. *Archives of Internal Medicine, 169*(13), 1233–1240. doi:10.1001/archinternmed.2009.177; Weller, S. & Davis, K. (2002). Condom effectiveness in reducing heterosexual HIV transmission. *The Cochrane Database of Systemic Reviews, 1.* doi:10.1002/14651858.CD003255; Sanchez, J., Campos, P., Courtois, B., Gutierrez, L., Carrillo, C., Alarcon, J., Gotuzzo, E., Hughes, J., Watts, D., Hillier, S. L., Buchanan, K., Holmes, K. K. (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. Retrieved from http://www.who.int/bulletin/volumes/82/6/454.pdf

26. See Endnote #3.

27. Teen Pregnancy Prevention Evidence Review. (n.d.). Retrieved from https://tppevidencereview.aspe.hhs.gov/EvidencePrograms.aspx; Farb, A. F., Margolis, A. L. (2016). The Teen Pregnancy Prevention Program (2010-2015): Synthesis of Impact Findings. *American Journal of Public Health, 106* (S1), S9-S15. doi: 10.2105/AJPH.2016.303367; Lugo-Gil, J., Lee, A., Vohra, D., Harding, J., Ochoa, L., Goesling, B., Mathematica Policy Research. (2018). Updated findings from the HHS Teen Pregnancy Prevention Evidence Review: August 2015 through October 2016. U.S. Department of Health and Human Services: Teen Pregnancy Prevention Evidence Review. Retrieved from https://tppevidencereview.aspe.hhs.gov/pdfs/Summary_of_findings_2016-2017.pdf; Abt and Associates. (2018). Reducing the Risk: Impact findings from the Teen Pregnancy Prevention Replication Study (Research Brief and Impact Evaluation Findings), November 5, 2018. U.S. Department of Health and Human Services. Retrieved from https://aspe.hhs.gov/pdf-report/reducing-risk-impacts-teen-pregnancy-prevention-replication-study-research-brief; Abt and Associates. (2018). *¡Cuídate!*: Impact findings from the Teen Pregnancy Prevention Replication Study (Research Brief and Impact Evaluation Findings), November 5, 2018. U.S. Department of Health and Human Services. Retrieved from https://aspe.hhs.gov/pdf-report/cuidate-impacts-teen-pregnancy-prevention-replication-study-research-brief

28. 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). 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. For a report of the study findings, see Chin H. B., Sipe, T. A., Elder, R., Mercer, S. L., Chattopadhyay, S., Jacob, V., Wethington, H. R., Kirby, D., Elliston, D. B., Griffith, M., Chuke, S. O., Briss, S. C., Ericksen, I., Galbraith, J. S., Herbst, J. H., Johnson, R. L., Kraft, J. M., Noar, S. M., Romero, L. M., Santelli, J., Community Preventive Services Task Force (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. doi: 10.1016/j.amepre.2011.11.006; 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. doi: 10.1016/j.amepre.2011.11.004; 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. doi: 10.1016/j.amepre.2011.11.003.

29. United Nations Educational, Scientific and Cultural Organization. (2009). *International Technical Guidance on Sexuality Education, Volume 1*, see pp.15-17. Retrieved from 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*. Retrieved from http://www.unaids.org/sites/default/files/media_asset/ITGSE_en.pdf

30. See Endnotes #11 and #14; 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. doi: 10.1016/j.amepre.2011.11.003.

31. Of the 79 discreet CSE interventions/programs analyzed in this database by 103 CSE studies, 12 of the programs are the subject of more than one study—eight in U.S. settings and four in international settings. (For one AE program in the U.S. there were two studies.) Because of this aspect of the research evidence, the quantitative results can be aggregated at one of two units or levels of analysis: using either the number of programs or the number of studies as the denominator when proportions are calculated. Arguments can be made for both methods, and we have chosen to report proportions calculated both ways where possible and appropriate. However, when summarizing *positive* program outcomes, we have tended to use the number of programs as the denominator. It is the smaller number, and so we err on the side of the type of program analyzed. Also, policymakers are typically choosing from a pool of programs not studies. When summarizing the prevalence of *negative* effects, we used the number of studies as the denominator—again, erring in favor of the program type and also providing the most accurate representation of the rate of negative outcomes.

32. These are listed by the study's first author, study date, and program name in Tables 5-7.

33. Philliber, A. E., Philliber, S., Brown, S. (2015). Evaluation of the Teen Outreach Program® in The Pacific Northwest; Accord, NY: Philliber Research & Evaluation. Retrieved from https://www.hhs.gov/ash/oah/sites/default/files/ash/oah/oah-initiatives/evaluation/grantee-led-evaluation/reports/ppgnw-final-report.pdf

34. Howard, M., & McCabe, J. (1990). Helping teenagers postpone sexual involvement. *Family Planning Perspectives, 22* (1), 21–26. doi: 10.2307/2135434; Aarons, S. J., Jenkins, R. R., Raine, T. R., et al. (2000). Postponing sexual intercourse among urban junior high school students – a randomized controlled evaluation. *Journal of Adolescent Health, 27*(4), 236–247. doi: https://doi.org/10.1016/S1054-139X(00)00102-6; Little, C. B., Rankin, A. (Unpublished). An evaluation of the Postponing Sexual Involvement curriculum. See page 127, item #35 in: https://powertodecide.org/sites/default/files/resources/primary-download/emerging-answers.pdf

35. 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. doi:10.1001/archpedi.160.8.772; 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. doi:10.2105/AJPH.2016.303371.

36. 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. doi.org/10.1037/0278-6133.21.2.177; Coyle, K. K., Basen-Enquist, K. M., Kirby, D. B., Parcel, G. S., Banspach, S. W., Collins, J. L., Baumler, E. R., Carvajal, S., Harris R. B. (2001). Safer Choices: Reducing Teen Pregnancy, HIV and STDs. *Public Health Reports, 1*(16), 82–93. doi.org/10.1093/phr/116.S1.82

37. Determining a rate of program effectiveness or success is limited by the number of studies that actually measured at least a 12-month post-program effect, as well as the high but unknown number of studies that did not make it into the database because of poor scientific quality. So the true success rate is unknown, and the estimates presented here should not be considered as absolute. They were computed to allow for quantitative comparisons across geographic region and program type.

38. Kelsey, M., Layzer, C., Layzer, J., Price, C., Juras, R., Blocklin, M., Mendez, J. (2016). Replicating ¡Cuídate!: 6-Month Impact Findings of a Randomized Controlled Trial. *American Journal of Public Health, 106*(S1), S70–S77. doi: 10.2105/AJPH.2016.303371; 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, doi.org/10.1016/j.jadohealth.2013.10.204; 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. doi:10.2105/AJPH.2016.303419; Philliber, A. E., Philliber, S., & Brown, S. (2015). Evaluation of the Teen Outreach Program® in The Pacific Northwest. Accord, NY: *Philliber Research & Evaluation*, retrieved from https://www.hhs.gov/ash/oah/sites/default/files/ash/oah/oah-initiatives/evaluation/grantee-led-evaluation/reports/ppgnw-final-report.pdf; Moberg, D., Piper, D. L. (1998). The Healthy for Life Project: Sexual risk behavior outcomes. *AIDS Education and Prevention 10*(2):128–48. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9573435; 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. For evidence of negative impact see also p. 53 in Laris, B. A., Kirby, D. (2007). One-Page Summaries of the Evaluations Referenced in Emerging Answers 2007. *National Campaign to Prevent Teen and Unplanned Pregnancy*); Abt and Associates. (2018). Reducing the Risk: Impact findings from the Teen Pregnancy Prevention Replication Study (Research Brief and Impact Evaluation Findings), November 5, 2018. U.S. Department of Health and Human Services. Retrieved from https://aspe.hhs.gov/pdf-report/reducing-risk-impacts-teen-pregnancy-prevention-replication-study-research-brief

39. 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. doi: https://doi.org/10.1111/j.1746-1561.2006.00135.x; 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. doi: https://doi.org/10.1177/0272431612449386 (Note: This was an evaluation of the first year of a three-year program that was a comprehensive sex education curriculum for the second and third years but an abstinence curriculum for the first year, with its effects evaluated 12 months later. See: https://tppevidencereview.aspe.hhs.gov/document.aspx?rid=3&sid=274&mid=2); 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. doi:10.1001/archpediatrics.2009.267; Piotrowski, Z., Hedeker, H., Hedeker, D. (2016). Evaluation of the *Be The Exception* Sixth-Grade program in rural communities to delay the onset of sexual behavior: Findings from an Innovative Teen Pregnancy Prevention Program. *American Journal of Public Health, 106*(S1), S132–S139. doi: 10.2105/AJPH.2016.303438; 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* (see also: Chin, et al., 2012 (in Endnote #27), p. 287, Figure 3 and p. 293, item #123. Retrieved from https://www.ajpmonline.org/article/S0749-3797(11)00906-8/fulltext); 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. Retrieved from https://www.institute-research.com/docs/Heritage_Keepers_Eval_(Weed,_Ericksen,_and_Birch,_2005).pdf; Weed, S. E., Ericksen, I. E., Lewis, A., Grant, G. E., Wibberly, K. H. (2008). An Abstinence Program's Impact on Cognitive Mediators and Sexual Initiation. *American Journal of Health Behavior, 32*(1), 60–73. doi: https://doi.org/10.5993/AJHB.32.1.6

40. See all references in Endnote #39, excluding Denny, et al., 2006, and Jemmott, et al., 2010.

41. In Endnote #39 see Denny, 2006; Jemmott, 2010; Petrowski, 2016.

42. Lieberman, L., Su, H. (2012). Impact of the Choosing the Best Program in Communities Committed to Abstinence Education. Sage Open, March 22, 2012. doi: 10.1177/2158244012442938.

43. Walker, E. M., Inoa, R., Coppola, N. (2016). Evaluation of Promoting Health Among Teens Abstinence-Only Intervention in Yonkers, NY. Sametric Research. Princeton, N.J. 08540. Retrieved from https://www.hhs.gov/ash/oah/sites/default/files/ash/oah/oah-initiatives/evaluation/grantee-led-evaluation/reports/program-reach-final-report.pdf; Farb, A. F., Margolis, A. L. (2016). The Teen Pregnancy Prevention Program (2010-2015): Synthesis of Impact Findings. *American Journal of Public Health*, 106 (S1), S9-S15. doi: 10.2105/AJPH.2016.303367.

44. 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. doi:10.5555/ajhb.2005.29.5.423; 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. doi:10.1001/jama.279.19.1529; 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. doi:10.1001/archpediatrics.2009.267; 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. doi: https://doi.org/10.1363/2910097; 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. doi:10.1016/j.jadohealth.2011.07.010; 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. doi: https://doi.org/10.1016/j.jadohealth.2013.10.204; Piotrowski, Z., Hedeker, H., Hedeker, D. (2016). Evaluation of the *Be The Exception* Sixth-Grade program in rural communities to delay the onset of sexual behavior: Findings from an Innovative Teen Pregnancy Prevention Program. *American Journal of Public Health, 106*(S1), S132–S139. doi: 10.2105/AJPH.2016.303438; 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. Retrieved from https://aspe.hhs.gov/report/impacts-four-title-v-section-510-abstinence-education-programs

45. 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. doi:10.1001/jama.279.19.1529.

46. These four studies, conducted by Mathematica Policy Research, Inc. and reported in a self-published paper (not an external peer-reviewed journal), have been cited by numerous reviewers as compelling evidence for AE failure. However, several features of these studies' designs raise doubts about their results. While these AE studies appear to have a strong experimental (randomized) evaluation design, in each case *the design was weakened in three different ways:* 1) *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 reduced sexual behavior in the treatment group, it would probably simultaneously diminish this behavior in the control group by reducing the number of sexual partners available to them. Thus, a reduction in sexual activity would likely occur in both the treatment and control groups as a result of the program, minimizing between group differences and therefore diminishing the size and statistical significance of the measured program effect. 2) *This design problem was compounded by another methodological factor—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 when measured 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 too long*—three to five years after the program (four to six years post-baseline). Because there was not additional abstinence message reinforcement during the long follow-up time interval, a post-program effect on behavior could not have reasonably been expected to persist at that point. Such unusually long follow-up times have rarely been employed in CSE studies and positive CSE effects that endure this long are almost non-existent (two out of 103 school-based CSE studies in the present database).These three factors in combination—unrealistically long follow-up time frames, randomizing treatment groups within schools, and extra-young subject populations—argue for viewing the findings of these four studies as "inconclusive" rather than as valid evidence of AE program failure. See: 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. Retrieved from https://aspe.hhs.gov/report/impacts-four-title-v-section-510-abstinence-education-programs

47. 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. doi:10.1016/j.jadohealth.2011.07.010; 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. doi: https://doi.org/10.1016/j.jadohealth.2013.10.204

48. Stephenson, J., Strange, V., Allen, E., Copas, A., Johnson, A., Bonell, C., Babiker, A., Oakley, A., the RIPPLE Study Team. (2008). The long-term effects of a peer-led sex education programme (RIPPLE): a cluster randomised trial in schools in England. *PLoS Medicine, 5*(11):1579–90. doi: https://doi.org/10.1371/journal.pmed.0050224

49. Okonofua, F. E., Coplan, P., Collins, S., Oronsaye, F., Ogunsakin, D., Ogonor, J. T., Kaufmanc, J. A., Heggenhougen, K. (2003). Impact of an intervention to improve treatment-seeking behavior and prevent sexually transmitted diseases among Nigerian youths. *International Journal of Infectious Diseases, 7,* 61–73. doi: https://doi.org/10.1016/S1201-9712(03)90044-0

50. Shuey, D. A., Babishangire, B. B., Omiat, S., Bagarukayo, H. (1999). Increased sexual abstinence among in-school adolescents as a result of school health education in Soroti district, Uganda. *Health Educ Res 14,* 411–419. doi: https://doi.org/10.1093/her/14.3.411

51. Magnani, R., MacIntyre, K., Karim, A. M., Brown, L., Hutchinson, P. (2005). The impact of life skills education on adolescent sexual risk behaviors in KwaZulu-Natal, South Africa. *Journal of Adolescent Health, 36,* 289–304. doi: https://doi.org/10.1016/j.jadohealth.2004.02.025

52. Diaz, M., Mello, M. B., Sousa, M. H., Cabral, F., Castro e Silva, R., Campos, M., Faúndes, A. (2005). Outcomes of three different models for sex education and citizenship programs concerning knowledge, attitudes, and behavior of Brazilian adolescents. *Cad Saude Publica, 21,* 589–597. Retrieved from https://www.scielosp.org/scielo.php?pid=S0102-311X2005000200026&script=sci_arttext; Diop, N. J., Bathidja, H., Touré, I. D., Dieng, T., Mané, B., RamaRao, S., Adamchak, S., Wong, E., Ndoye, C. A., Sy, A., Fall, B. (2004). Improving the reproductive health of adolescents in Senegal. *Frontiers in Reproductive Health.* New York: Population Council; 2004. Retrieved from https://www.researchgate.net/profile/Nafissatou_Diop2/publication/241175190_Improving_the_Reproductive_Health_of_Adolescents_in_Senegal/links/53f36f770cf2da8797447dc4/Improving-the-Reproductive-Health-of-Adolescents-in-Senegal.pdf; Dupas, P. (2011). Do teenagers respond to HIV risk information? Evidence from a field experiment in Kenya. *American Economic Journal: Applied Economics, 3*(1), 1–34. doi=10.1257/app.3.1.1; Fitzgerald A. M., Stanton, B. F., Terreri, N., Shipena, H., Li, X., et al. (1999). Use of Western-based HIV risk-reduction interventions targeting adolescents in an African setting. *Journal of Adolescent Health, 25,* 52–61. doi: https://doi.org/10.1016/S1054-139X(98)00120-7; Jewkes, R., Nduna, M., Levin, J., Jama, N., Dunkle, K., Puren, A., Duvvury, N. (2008). Impact of stepping stones on incidence of HIV and HSV-2 and sexual behaviour in rural South Africa: Cluster randomized controlled trial. *BMJ,337*, a506. doi: https://doi.org/10.1136/bmj.a506; Merakou, K., Kourea-Kremastinou, J. (2006). Peer education in HIV prevention: An evaluation in schools. *European Journal of Public Health, 16,* 128–32. doi:10.1093/eurpub/cki162; Ross, D. A., Changalucha, J., Obasi, A. I., Todd, J., Plummer, M. L., Cleophas-Mazigec, B., Anemonaa, A., Everetta, D., Weissa, H. A., Mabeya, D. C., Grosskurtha, H., Hayeset, R. J. (2007). Biological and behavioural impact of an adolescent sexual health intervention in Tanzania: a community-randomized trial. *AIDS 21,* 1943–1955. doi: 10.1097/QAD.0b013e3282ed3cf5; Smith, E. A., Palen, L. A., Caldwell, L. L., Flisher, A. J., Graham, J. W., Mathews, C., Wegner, L., Vergnani, T. (2008). Substance use and sexual risk prevention in Cape Town, South Africa: an evaluation of the HealthWise program. *Prevention Science, 9,* 311–321. doi: 10.1007/s11121-008-0103-z; Visser, M. (2007). HIV/AIDS prevention through peer education and support in secondary schools in South Africa, SAHARA-J: *Journal of Social Aspects of HIV/AIDS, 4*:(3), 678-694. doi: 10.1080/17290376.2007.9724891.

53. In Endnote #52, see Diop, et al., 2004, Dupas, 2011, and Visser, 2007.

54. See Endnote #49.

55. See Endnote #50.

56. See references in Endnote #52, excluding Dias, et al., 2005, and Merakou, et al., 2006.

57. See references in Endnotes #34, #36, #48, #49, and #50.

58. See reference in Endnote #48.

59. See references in Endnotes #38 and #52.

60. Determining a rate of program effectiveness or success is limited by the number of studies that actually measured at least a 12-month post-program effect, as well as the high but unknown number of studies that did not make it into the database because of poor scientific quality. So the true success rate is unknown, and the estimates presented here should not be considered as absolute. They were computed to allow for quantitative comparisons across geographic region and program type. On the other hand, every study in the database had the opportunity to detect and report negative effects, so these estimates might be considered a more true estimation of a *minimum* incidence of negative program impact.

61. United Nations Educational, Scientific and Cultural Organization. (2018). *International Technical Guidance on Sexuality Education: An Evidence-Informed Approach,* p. 12, 28. Retrieved from http://www.unaids.org/sites/default/files/media_asset/ITGSE_en.pdf

62. 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. doi: https://doi.org/10.1016/j.amepre.2011.11.003

63. Teen Pregnancy Prevention Evidence Review. (n.d.). Retrieved from https://tppevidencereview.aspe.hhs.gov/EvidencePrograms.aspx

64. We did not have enough information about program and study authorship to look at this ratio for the international studies.

65. For example, see "Standards for Effectiveness: 5f" in 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),

66. United Nations Educational, Scientific and Cultural Organization. (2018). *International Technical Guidance on Sexuality Education: An Evidence-Informed Approach,* p. 29. Retrieved from http://www.unaids.org/sites/default/files/media_asset/ITGSE_en.pdf

67. United Nations Educational, Scientific and Cultural Organization. (2018). *International Technical Guidance on Sexuality Education: An Evidence-Informed Approach,* p. 18, 29. Retrieved from http://www.unaids.org/sites/default/files/media_asset/ITGSE_en.pdf

68. These were: *Be Proud Be Responsible, Positive Prevention Plus, Promoting Health Among Teens/Comprehensive, Reducing the Risk,* and *Teen Outreach Program*, listed at: https://www.blueprintsprograms.org/programs. Unfortunately, two recent independent evaluation studies funded by the U.S. federal TPP program have found harmful effects by the latter two programs, which would disqualify them from the *Blueprints* designation as "Promising" programs (see Endnotes #18 and #38).

69. Blueprints Programs. Retrieved from https://www.blueprintsprograms.org/programs and https://www.blueprintsprograms.org/faq#what-difference-mp

Table 1. School-Based Sex Education: Number of Studies Finding EVIDENCE OF EFFECTIVENESS

	School-based Comprehensive Sex Education 103 Studies				**Abstinence-only Education 17 Studies**
***Effectiveness Criteria*:** *a protective effect for the intended population on sexual initiation, condom use, pregnancy, or STDs, at least 12 months post-program, without other negative outcomes*	**Combined Non-U.S. & U.S.** 103 studies (79 programs)	**Non-U.S.** 43 studies (39 programs)	**Africa** (Subset of non-U.S.) 29 studies (26 programs)	**U.S.** 60 studies (40 programs)	**U.S.** 17 studies (16 programs)
Reduced Pregnancy	1	1	0	0	0
Reduced STDs	1	1	1	0	0
Increased Abstinence *(Delayed Sexual Initiation)*	2	1	1	1	7
Increased Consistent Condom Use *(CCU)*	0	0	0	0	0
Increased Condom Use Frequency or Use at Last Sex when CCU was not measured	2	0	0	2	0
Dual Benefit *(Increased Abstinence & Condom Use in the same population)*	0	0	0	0	0
Total # of Studies with Evidence of Effectiveness	**6**	**3**	**2**	**3**	**7**
Independent Evidence # of independent studies (not by the program's developers) that found evidence of effectiveness	**1**	**1**	**0**	**0**	**5**

Table 2. School-Based Sex Education: EVIDENCE of PROGRAM SUCCESS vs. FAILURE					
	School-Based Comprehensive Sex Education 103 Studies				**Abstinence-only Education 17 Studies**
Effectiveness Criteria: *a protective effect for the intended population on sexual initiation, condom use, pregnancy, or STDs, at least 12 months post-program, without other negative outcomes*	**Combined Non-U.S. & U.S.** 103 studies (79 programs)	**Non-U.S.** 43 studies (39 programs)	**Africa** **(Subset of non-U.S.)** 29 studies (26 programs)	**U.S.** 60 studies (40 programs)	**U.S.** 17 studies (16 programs)
Program Success or Effectiveness[a] # of programs finding evidence of effective-ness as a proportion of the # of programs that measured effectiveness	6[b]/47 13%	3/27 11%	2/19 11%	3/20 15%	7/15 47%
Program Failure % of programs that measured effectiveness and did not find it	**87%**	**89%**	**89%**	**85%**	**53%**

[a] It should be noted that estimating a rate of effectiveness/success is limited by the number of studies that actually measured at least a 12-month post-program effect, as well as the high number of studies that did not make it into the database because of poor scientific quality. So the true incidence of program success is unknown and these estimates should not be considered absolute but only as representing the evidence available in this database.

[b] Of the 6 studies, 5 were conducted by the program developers, leaving one that provided independent evidence of effectiveness.

Table 3. School-Based Sex Education: EVIDENCE of HARMFUL EFFECTS					
	School-Based Comprehensive Sex Education 103 Studies				Abstinence-only Education 17 Studies
Negative Effects a worsening of sexual health or risk behaviors for the intended population or a substantial subgroup, for any duration	**Combined Non-U.S. & U.S.** 103 studies (79 programs)	**Non-U.S.** 43 studies (39 programs)	**Africa** (Subset of non-US) 29 studies (26 programs)	**U.S.** 60 studies (40 programs)	**U.S.** 17 studies (16 programs)
Increased Pregnancy	1	0	0	1	0
Increased STDs	1	1	1	0	0
Increased Sexual Activity *(Initiation/Frequent/Recent Sex)*	9	5	3	4	0
Decreased Condom Use	3	1	1	2	0
Increased Oral Sex	2	0	0	2	0
Increased #Sex Partners	3	2	2	1	1
Increase in Forced or Coerced Sex	2	2	2	0	0
Increase in Paid Sex	1	1	1	0	0
Total #Negative Effects	22	12	10	10	1
Net #of Studies and Programs with Negative Effects (for some studies or programs there was more than one harmful effect)	**16 studies 16%**	**9 studies 21%**	**7 studies 24%**	**7 studies 12%**	**1 study 6%**
	15 programs 19%	**9 programs 23%**	**7 programs 27%**	**6 programs 15%**	**1 program 6%**

Table 4. School-Based Sex Education: EVIDENCE of EFFECTIVENESS vs. HARM					
	School-Based Comprehensive Sex Education 103 Studies				Abstinence-only Education 17 Studies
Effectiveness Criteria: *a protective effect for the intended population on sexual initiation, condom use, pregnancy, or STDs, at least 12 months post-program, without other negative outcomes*	**Combined Non-U.S. + U.S.** 103 studies (79 programs)	**Non-U.S.** 43 studies (39 programs)	**Africa** (Subset of non-U.S.) 26 programs (29 studies)	**U.S.** 60 studies (40 programs)	**U.S.** 17 studies (16 programs)
Effectiveness (Success) # of studies that found evidence of effectiveness	6 studies	3 studies	2 studies	3 studies	7 studies
Negative Effects (Harm) # of studies that found a worsening of sexual health or risk behavior for the intended population or a major sub-group, lasting any duration	16 studies	9 studies	7 studies	7 studies	1 study

Table 5. U.S. School-based Comprehensive Sex Education (CSE): 60 Studies of 40 Programs.

PROGRAM AND STUDY			STUDY OUTCOMES									
PROGRAM NAME	STUDY 1st AUTHOR & YEAR	Independent Evaluator?	Negative Effect	Post-Program Main Effect on Most-Protective Indicators				Post-Program Main Effect on Less-Protective Indicators				Dual Benefit
				Sexual Initiation	Consistent Condom Use	Pregnancy	STDs	CondomFrequency	Frequ/RecentSex	UnprotectedSex	#SexPartners	12mo. Post-Program
Aban Aya (Classroom Version Only)	Flay, 2004	No	No	NM	NM	NM	NM	NS	NS	NM	NM	NM
AIDS Prevention Program	Siegel, 1995	?	No	NM	NM	NM	NM	NS	NM	NM	NS	NM
All4You	Coyle, 2006	No	No	NS	NM	NS	NM	6 months Only	6 months Only	6 months Only	NS	NS
All4You	Coyle, 2013	No	No	NS	NM	NM	NM	NS	NM	NS	NM	NS
All4You2 (Classroom Version Only)	Coyle, 2013	No	No	NS	NM	NM	NM	NS	NM	4 months Only	NM	NS
Be Proud Be Responsible (held on Saturday)	Jemmott, 1999	No	No	NM	NM	NM	NM	NM	NS	6 months	NS	NM
Be Proud Be Responsible-(held on School Day)	Borawski, 2009	Yes	No	NS	NS	NM	NM	NM	NS	NS	NM	NM
[Blake] HIV/STD Prevention Curriculum	Blake, 2000 (unpub.)	Yes	No	NS	NS	NM	NM	NM	NS	NM	NM	NM
[Boyer] HIV/STD Prevention Curriculum	Boyer, 1997	No	No	NM	NM	NM	NM	NS	NM	NM	NS	NM
Crossroads Program (adaptation BPBR)	Slater & Mitschke,2015	?	No	NM	NM	NS	NM	NM	NM	6 months Only	NM	NM
¡Cuídate! (held on Saturday)	Villaruel, 2006	No	No	NS	12 months	NM	NM	NS	12 months	12 months	12 months	NS
¡Cuídate! (held on School Day)	Kelsey,2016;AbtAssoc,2015	Yes	3 Negative Effects	Negative Effect	NM	NS	NS	NS	Negative Effect	Negative Effect	NM	NM
Draw the Line,Respect the Line	Coyle, 2004	No	No	12 mo-Subgroup O	NM	NM	NM	NS	NS	NM	NS	NS
Focus on Kids/West Virginia	Stanton, 2005	No	No	NM	NM	NM	NM	NS	NS	NM	NM	NM
Gender Matters	Smith, Kim, et al, 2015	No	No	NS	NM	NM	NM	NM	NS	NS	NM	NM
Get Real About AIDS	Main, 1994	?	No	NS	NM	NM	NM	6 months	NS	NM	6 months	NM
Get Real - 7th & 8th Grade Only	Grossman, 2014	?	No	<9 months	NM	NM	NM	NM	NM	NM	NM	NM
H.A.R.T. (adaptation of B.A.R.T.)	Boston Medical Center	?	No	NS	NM	NM	NM	NM	NS	NS	NM	NM
Health Teacher	MathmaticaPolicyResearch	No	No	NS	NM	NM	NM	NM	NM	NM	NM	NM
Healthy & Alive!	Middlestadt, [Unpubl]	No	No	NS	NS	NM	NM	NM	NS	NM	No	NM
Healthy for Life-Version 1 (Age-based)	Moberg, 1998/2000	?	No	NS	NS	NM	NM	NM	NS	NM	NM	NS
Healthy for Life-Version 2 (7th Gr.Intensive)	Moberg, 1998/2000	?	Recent Sex	NS	NS	NM	NM	NM	Negative Effect	NM	NM	NS
Healthy Oakland Teens	Ekstrand, 1996(AIDSConf)	?	No	8 to 11 months	NM	NM	NM	NM	NM	NM	NM	NM
HIV Prevention Interventions	Fisher, 2002	No	No	NS	NM	NM	NM	12 months	NM	NM	NM	NS
It's Your Game: Keep It Real/Risk Reduction	Tortolero, 2010	No	No	12 months	NM	NM	NM	NS	12 months	NM	NS	NS
IYG - Risk Reduction	Markham, 2012/2014	No	#Sex Partners	10 months Only	10 months Only	NM	NM	NM	10 months O	10 months O	Negative Effect	NS
IYG - Risk Reduction (South Carolina)	Potter, 2016a	Yes	Sexual Initiation	Negative Effect	NM	NM	NM	NM	NS	NS	NM	NS
IYG - Risk Reduction (Texas)	Coyle, 2016b	Yes	No	NS	NM	NM	NM	NM	NM	NM	NM	NM
Making Proud Choices! (held on Saturday)	Jemmott, 1998	No	No	NS	3 months Only	NM	NM	3, 6, 12 months	12mo-Subgroup	12mo-Subgroup	NM	NS
Need To Know	UofTxHlthSciCtrSanAntonio	?	No	NS	NM	NM	NM	NM	NM	NM	NM	NM
Positive Prevention	LaChausse, 2006	Yes	No	6 months	NM	NM	NM	NS	6 months	NM	NM	NM
Positive Prevention PLUS	LaChausse, 2015/2016	Yes	No	6 months	NM	NS	NM	NM	NM	6 months	NM	NM
Postponing Sexual Involvement(PSI)	Howard&McCabe,1990	No	No	12 months	NM	NM	NM	NM	NS	NM	NM	NM
Postponing Sexual Involvement	Aarons, 2000	Yes	No	NS	NM	NM	NM	NM	NM	NM	NM	NM
PSI & HumanSexuality-adapted	Little & Rankin, unpub.	?	No	NS	NM	NM	NM	NM	NS	NM	NS	NM
Project IMPPACT Inwood House	Lieberman, 2000	No	No	NS	NM	NS	NM	NS	NM	NM	NM	NS
Project LIGHT	Lightfoot, 2007	No	No	NM	NM	NM	NM	NS	3 months	NS	3 months	NM
Project SNAPP	Kirby, 1997	Yes	Contraception	NS	NM	NS	NS	Negative Effect	NS	NM	NM	NS
PromotingHealthAmongTeens!/CSE (on Saturday)	Jemmott, 2010	No	No	NS	NS	NM	NM	NM	NS	NS	24 months	NS
Reach for Health	O'Donnell, 1999	?	No	NS	NM	NM	NM	NM	NS	NM	NM	NM
Reducing the Risk (RTR)	Kirby, 1991	No	No	NS	NM	NS	NM	NS	NM	NS	NM	NM
RTR	Barth, 1992	No	No	NS	NM	NS	NM	NS	NS	NM	NM	NM
RTR	Hubbard, 1998	Yes	No	18months	NM	NM	NM	NM	NM	NM	NM	NM
RTR	Kelsey,2016;AbtAssoc,2018	Yes	Recent/OralSex	NS	NM	24mo-Subgroup O	NS	NM	Negative Effect	NS	NM	NS
RTR	Zimmerman, 2008a	No	No	NS	NM	NM	NM	NS	NM	NM	NM	NM
Reducing the Risk (RTR)-modified1	Zimmerman, 2008b	No	No	NS	NM	NM	NM	NS	NM	NM	NM	NM
RTR	Reyna & Mills, 2014a	Yes	No	NS	NM	NM	NM	NS	NM	NS	NS	NS
Reducing the Risk (RTR)-modified2	Reyna & Mills, 2014b	No	No	12 months	NM	NM	NM	NS	NM	NS	12 months	NS
Rochester AIDS Prevention Project(RAPP)	Siegel, 2001	?	No	NS	NM	NM	NM	NM	NM	NM	NM	NM
Rochester AIDS Prevention Project(RAPP)	Aten, 2002	?	No	NS	NM	NM	NM	NM	NM	NM	NM	NM
Safer Choices	Coyle, 2001	No	No	NS	NM	NM	NM	12 months	NS	NM	NS	NS
Teen Outreach Program (TOP)	Allen, 1997	Yes	No	NM	NM	atProgEnd-SubgrpO	NM	NM	NM	NM	NM	NM
TOP (Florida)	Daley, 2015	Yes	No	at Prog. End O	NM	at Prog.End O	NM	NM	NM	NM	NM	NM
TOP (Hennepin, MN)	Francis, 2015	Yes	No	NS	NM	NM	NM	NM	NS	NS	NM	NM
TOP (Chicago)	Seshadri, 2015	Yes	No	NM	NM	NM	NM	NM	NS	NS	NM	NM
TOP (GNWPP)	Philliber, 2016	Yes	Pregnancy	NM	NM	Negative Effect	NM	NM	NS	NM	NM	NM
TOP (held on Saturday)	Robinson, 2016	Yes	No	NS	NM	NM	NM	NM	NM	NS	NM	NM
[Walter&Vaughn]AIDS Prevention	Walter & Vaughn, 1993	?	No	NS	3 months	NM	NS	NM	NM	NM	3 months	NM
Wise Guys	Gottsegen, 2001	No	No	NS	NM	NM	NM	NM	NM	NM	NM	NM
Youth AIDS Prevention Project (YAPP)	Levy, 1995	No	No	NS	NM	NM	NM	NS	NS	NM	NS	NS

NM = Did not measure this outcome; NS = The study measured this outcome but the effect was not statistically significant at p<.05; F = Females; M = Males; O = Only; at Prog. End= measured at the program's endpoint

Green = Evidence of Program Effectiveness: A significant effect on a key protective indicator, at least 12 months post-program, on the intended target population (not just a subgroup), without other negative effects.

Blue = Evidence of Program Potential; Brown = Evidence of Program Failure (measured this outcome but failed to find a significant effect); Red = Evidence of Negative Program Impact; Grey = Program produced both positive and negative effects

Table 6. U.S. School-based Abstinence Education (AE): 17 Studies of 16 Programs.

PROGRAM AND STUDY			STUDY OUTCOMES							
			Post-Program Main Effect on Most-Protective Indicators				Post-Program Main Effect on Less-Protective Indicators			Dual Benefit
PROGRAM NAME	STUDY 1st AUTHOR & YEAR	Independent Evaluator?	Sexual Initiation	Consistent Condom Use	Pregnancy	STDs	CondomFrequency	Frequent/RecentSex	# Sex Partners	12mo. Post-Program
Choosing the Best	Weed, 2008	Yes	12 months	NM	NM	NM	NM	NM	NM	NM
Choosing the Best	Lieberman, 2012	Yes	3 months only	NM	NM	NM	NM	NS	NM	NM
For Keeps	Borawski, 2005	Yes	NS	NS	NM	NM	NM	NM	5 months	NM
Get Real-6th Grade Only	Erkut, 2012	Yes	12 months	NM	NM	NM	NM	NM	NM	NM
Heritage Keepers	Weed, 2011	Yes	12 months	NM	NM	NM	NM	NM	NM	NM
It's Your Game: Keep It Real/Abstinence-Only	Markham, 2012/2014	No	NS	(Unprotected Sex) 10/24mo	NM	NM	NM	NM	Negative Effect	NS
Making a Difference! An Abstinence Program	Jemmott, 1998	No	3 months only	NS	NM	NM	12 months	NM	NM	NS
My Choice, My Future	Trenholm, 2007	Yes	NS	NM	NS	NS	NS	NS	NS	NS
Positive Potential - 6th Grade	Piotrowski, 2016	Yes	12 months	NS	NM	NM	NM	12 months	NM	NS
Postponing Sexual Involvement/Abstinence-Only	Kirby, 1997	Yes	NS	NM	NS	NS	NS	NM	NS	NS
Promoting Health Among Teens/Abstinence-Only	Jemmott, 2010	No	24 months	NS	NM	NM	NM	24 months	NM	NS
Reasons of the Heart	Weed, 2008	Yes	12 months	NM	NM	NM	NM	NM	NM	NM
ReCapturing the Vision	Trenholm, 2007	Yes	NS	NM	NS	NS	NS	NS	NS	NS
Sex Can Wait	Denny, 2006	No	18 months	NM	NM	NM	NM	18 months	NM	NM
Teens In Control	Trenholm, 2007	Yes	NS	NM	NS	NS	NS	NS	NS	NS
Sex Respect, Teen-Aid, Values and Choices	Weed, 1992	Yes	NS	NM	NM	NM	NM	NM	NM	NM
WAIT Training	Rue, 2005	Yes	NS	NM	NM	NM	NM	NM	NM	NM

NM = Did not measure this outcome; NS = The study measured this outcome but the effect was not statistically significant at p<.05

F = Females; M = Males; O = Only; at Prog. End = measured at the program's endpoint

Green = Evidence of Program Effectiveness: A significant effect on a key protective indicator, at least 12 months post-program, on the intended target population (not just a subgroup), without other negative effects.

Blue = Evidence of Program Potential; Brown = Evidence of Program Failure (measured the outcome but the effect was not significant); Red = Evidence of Negative Impact; Grey = Program produced both positive and negative effects

Table 7. International (non-U.S.) School-Based Comprehensive Sex Education (CSE): 43 Studies of 39 Programs.

STUDY & PROGRAM			STUDY OUTCOMES								
				Post-Program Main Effect on Most-Protective Indicators				Post-Program Main Effect on Less-Protective Indicators				Dual Benefit
Study 1st Author & Year*	Program Name	Country	Negative Effects	Sexual Initiation	Consistent Condom Use	Pregnancy	STDs	Any Condom Use	Frequent/RecentSex	Unprotected Sex	# Sex Partners	12mo. Post-Program
Ajuwon, 2007	Reproductive Health Education	Nigeria, Africa	No	NM	NM	NM	NM	End Of Program	NS	NM	NM	NM
Aderibigbe, 2008	HealthEducationOnRiskyBehavior	Nigeria, Africa	No	NM	NM	NM	NM	3 months	NS	NM	3 months	NM
Agha, 2004	Peer-led HIV/AIDS Prevention	Zambia, Africa	No	NM	NM	NM	NM	NS	6 months	NM	6 months	NM
Borgia, 2005	Peer-led HIV/AIDS Prevention	Rome, Italy	No	NM	NM	NM	NM	NS	NM	NM	NS	NM
Cartagena, 2006	Peer-led HIV/AIDS Prevention	Mongolia	No	NM	3yr-PostBaseline-Subgroup O	NM	NM	NM	NM	NM	NM	NM
Daboer, 2008	HIV/AIDS Health Education	Nigeria, Africa	No	6 months	NM	NM	NM	NM	NS	NM	NM	NM
Dente, 2005	HIV Education & Counseling	Uganda, Africa	No	NS	NS	NM	NM	NM	NM	NM	NS	NS
Diaz, 2005 (1)	EDUCARTE	Rio DeJaneiro, Brazil	No	NM	NM	NM	NM	NS	NS	NM	NM	NM
Diaz, 2005 (2)	Education: An Exercise In Citizenship	Salvador, Brazil	No	NM	NM	NM	NM	NS	NS	NM	NM	NM
Diaz, 2005 (3)	Sexuality & Affectivity Education	Belo Horizonte, Brazil	Recent Sex	NM	NM	NM	NM	NS	Negative Effect	NM	NM	NM
Diop, 2004	Accompanying The Future	Senegal, Africa	SI & Coerced Sex	Negative Effect	NM	NM	NM	NM	NS	15mo PostBaseline	NS	NS
Doyle, 2010	MEMA kwa Vijana	Tanzania, Africa	No	NS	NM	NS	NS	NS	NS	NM	NS	NS
Duflo, 2006	Critical Thinking	Kenya, Africa	No	NS	NM	NM	NM	6mo Subgroup O	NM	NM	NS	NM
Duflo, 2015	Critical Thinking	Kenya, Africa	No	NM	NM	NS	NS	NM	NM	NM	NM	NM
Dupas, 2011	Relative Risk Info Campaign	Kenya, Africa	SI & #Partners	Negative Effect	NM	1 year	NM	NS	NM	NM	Negative Effect	NS
Fawole, 1999	Sch-Bsd HIV/AIDS Education Prog.	Nigeria, Africa	No	NM	NS	NM	NS	NS	6 months	NM	6 months	NM
Fitzgerald, 1999	My Future is My Choice(MFMC)	Namibia, Africa	Condom Use	NS	NM	NM	NM	Negative Effect	NS	NM	NS	NM
Harvey, 2000	DramAide	South Africa	No	NS	NM	NM	NS	6 months	NM	NM	NS	NM
Henderson, 2007	SHARE	Scotland, UK	No	NM	NM	NS	NM	NM	NM	NM	NM	NM
James, 2006	HIV LifeSkills Program	KZN, SoAfrica	No	4mo Subgroup O	NM	NM	NM	atProgEnd-Subgroup O	atProgEnd-Subgroup O	NM	NM	NM
Jemmott, 2015	Let Us Protect Our Future	South Africa	No	NS	NS	NM	42mo Subgroup O	NS	NS	12 months	NS	NS
Jewkes, 2008	SteppingStones	South Africa	Paid Sex	NM	NM	NS	12 months	NS	Negative Effect	NM	NS	NS
Karnell, 2006	Our Times Our Choices(RTRadapted)	South Africa	No	NS	NM	NM	NM	NS	NM	NM	NM	NM
Li, 2008	FocusOnKids adapted	Nanjing,China	No	NS	NM	NM	NM	NS	NM	NM	NM	NM
Magnini, 2005	HIV LifeSkills Program	South Africa	No	NS	NS	NM	NM	2yr Post-Baseline	NS	NM	NS	NS
Martinez-Donate, 2004	CSE+Condom Distribution	Tijuana, Mexico	No	6 months	NM	NM	NM	NS	NS	NS	NM	NM
Maticka-Tyndale , 2010	PSABH	Kenya, Africa	No	18mo Subgroup O	NM	NM	NM	30mo Subgroup O	18mo(in past 3mo)	NM	NM	NS
Matthews, 2010 - Site 1	SATZ HIV Prevention Program	Cape Town, So.Africa	No	NS	NM	NM	NM	NS	NM	NM	NM	NS
Matthews, 2010 - Site 2	SATZ HIV Prevention Program	Mankweng, So.Africa	No	NS	NM	NM	NM	NS	NM	NM	NM	NS
Matthews, 2010 - Site 3	SATZ HIV Prevention Program	DarEsSalaam,Tanzania	No	12mo Subgroup O	NM	NM	NM	NS	NM	NM	NM	NS
Merakou, 2006	Peer-led HIV/AIDS Prevention	Athens, Greece	Sexual Initiation	Negative Effect	NM	NM	NM	6mo Post-Baseline	NM	NM	NM	NM
Okonofua, 2003	WomensHealth&ActionResourceCtr.	Nigeria, Africa	No	NM	NM	NM	1yrPostBsline	1yrPostBsl-Subgroup O	NM	NM	NM	NM
Ross, 2007	MEMA kwa Vijana	Tanzania, Africa	STDs	NS	NM	NS	Negative Effect	3yr Post-Baseline	NM	NM	3yrPostBsl-Subgroup O	NS
Smith, 2008	HealthWise South Africa	South Africa	Sexual Initiation	Negative Effect	NS	NM	NM	NM	NS	NM	NM	NM
Stanton, 1998	My Future is My Choice(MFMC)	Namibia, Africa	No	12mo Subgroup O	NM	NM	NM	6mo Subgroup O	NS	NM	NS	NS
Shuey, 1999	School Health Education	Uganda, Africa	No	2yr Post-Baseline	NM	NM	NM	NM	NS	NM	NS	NS
Stephenson, 2008	RIPPLE	England, UK	No	NS	NM	54 months	NM	NS	NM	NS	NM	NS
Taylor, 2014	Teen Pregnancy Prevention Prog.	KZN, South Africa	No	NS	NS	NS	NM	5 months	NM	NM	NM	NM
Thato, 2008	BeforeYouKnow/SmartLove	Thailand	No	NM	NS	NM	NM	NM	6 months	NM	NM	NM
Visser, 2007	Peer-led HIV/AIDS Prevention	South Africa	ForcedSex&#Partners	18mo Post-Baseline	NS	NM	NM	NM	18mo Post-Baseline	NM	Negative Effect	NS
Walker, 2006	HIVPrevention+EmergencyContrcptn	Mexico	No	NM	NM	NM	NM	NS	NM	NM	NM	NM
Wight, 2002	SHARE-Interim Evaluation	Scotland, UK	No	NS	NM	NS	NM	NS	NM	NS	NM	NM
Ye, 2009	HIV Education	China	No	NM	NM	NM	NM	NS	NM	NM	NM	NM

*In this table the studies were ordered alphabetically by author instead of by program because many of these international CSE programs had similar (generic) names, making the author's name a better way to identify studies/programs.

NM = Did not measure this outcome; NS = The study measured this outcome but the effect was not statistically significant at p<.05; F = Females; M = Males; O = Only; at Prog. End= measured at the program's endpoint; SI = Sexual Initiation

Green = Evidence of Program Effectiveness: A significant effect on a key protective indicator, at least 12 months post-program, on the intended target population (not just a subgroup), without other negative effects.

Blue = Evidence of Program Potential; Brown = Evidence of Program Failure (measured this outcome but failed to find a significant effect); Red = Evidence of Negative Program Impact; Grey = Program produced both positive and negative effects

Note: Follow-up time periods shown in the cells indicate duration of effect *after the program's end* unless labeled as "Post-Baseline" which means effects were measured immediately following the end of a long-term program.