Abstract

This paper reviews 83 studies that measure the impact of curriculum-based sex and HIV education programs on sexual behavior and mediating factors among youth under 25 years anywhere in the world. Two thirds of the programs significantly improved one or more sexual behaviors. The evidence is strong that programs do not hasten or increase sexual behavior but, instead, some programs delay or decrease sexual behaviors or increase condom or contraceptive use. Effective curricula commonly incorporated 17 characteristics that describe the curricula development; the goals, objectives, and teaching strategies of the curricula themselves; and their implementation. Programs were effective across a wide variety of countries, cultures, and groups of youth. Replications of studies also indicate that programs remain effective when implemented by others in different communities, provided all the activities are implemented as intended in similar settings. © 2007 Society for Adolescent Medicine. All rights reserved.

Keywords: Adolescent; Sex education; HIV education; Prevention; Pregnancy in adolescence; Sexually transmitted diseases; HIV; Sexual behavior; Contraceptive behavior; Randomized controlled trials

Despite substantial declines in teen pregnancy in the United States, pregnancy and birth rates are still very high [1], as are the rates of several sexually transmitted diseases (STDs) [2]. In other developed countries, teen pregnancy rates are typically much lower than in the United States, but some STD rates are increasing and becoming serious problems [3,4]. In many developing countries, as increasing numbers of young people delay marriage until they are older, they have also become more likely to have sex before marriage; thus premarital pregnancy rates and some STD rates are increasing [5].

Throughout the world, many people view sex and HIV/STD education programs as a partial solution to these problems [6]. Indeed, sex and HIV/STD education programs that are based on a written curriculum and that are implemented among groups of youth in school, clinic, or community settings are a promising type of intervention to reduce adolescent sexual risk behaviors. They are often well-designed to be implemented in schools, where they can potentially reach large numbers of youth, yet they can also be implemented in clinic and community settings where they can also reach other youth, including potentially higher risk youth who have dropped out of school.

There have been many previous reviews of sex and HIV education programs’ impact on behavior, but they typically have been limited to a particular geographical area such as the United States [7,8] or sub-Saharan Africa [9], or they are now somewhat dated [7,10], or they do not analyze in depth the characteristics of effective programs.

This review attempts to overcome those and other limitations. It includes 83 studies from all countries, both developed and developing; it is more recent; it includes programs for youth up to age 25; it presents results for mediating factors as well as outcome behaviors; and it examines in depth the characteristics of effective programs that changed behavior.

This paper addresses two questions: What are the effects, if any, of curriculum-based sex and HIV education programs on sexual risk behaviors, STD and pregnancy rates, and mediating factors such as knowledge and attitudes that
affect those behaviors? What are the common characteristics of the curriculum-based programs that were effective in changing sexual risk behaviors?

Methods

Identification of evaluation studies

To be included in this review, each study had to meet the following criteria:

The program had to:

- Be a curriculum- and group-based sex or HIV education program (as opposed to an intervention involving only spontaneous discussion, only one-on-one interaction, or only broad school, community or media awareness activities).
- Focus on adolescents or young adults ages 9 to 24 years.

The research methods had to:

- Include a reasonably strong experimental or quasi-experimental design with both intervention and comparison groups and both pretest and posttest data collection.
- Have a sample size of at least 100.
- Measure program impact upon one or more of the following sexual behaviors: initiation of sex, frequency of sex, or number of sexual partners; use of condoms or contraception more generally; composite measures of sexual risk (e.g., frequency of unprotected sex); pregnancy rates; birth rates; and STD rates.
- Measure impact on those behaviors that can change quickly (i.e., frequency of sex, number of sexual partners, use of condoms or contraception or sexual risk-taking) for at least 3 months, or measure impact on those behaviors or outcomes that change less quickly (i.e., initiation of sex, pregnancy rates, or STD rates) for at least 6 months.

The study had to be completed or published in 1990 or thereafter. (To be as inclusive as possible, studies did not have to be published in peer-reviewed journals to be included in the review.)

Studies meeting these criteria were identified in several ways, including: searches of 10 databases (PubMed, PsychInfo, Popline, Sociological Abstracts, Psychological Abstracts, Bireme, Dissertation Abstracts, ERIC, CHID, and Biologic Abstracts), reviews of past issues of 12 journals; contacts with researchers at professional meetings and those in the process of completing studies; a review of reports, training materials, and process evaluation reports; and previous literature searches and reviews from various authors.

Analysis of study results

All identified studies meeting these criteria were reviewed and specific information from each study was summarized in a one-page summary that included key data such as characteristics of the sample, the characteristics of the intervention, research methods, and effects on both sexual risk behaviors and mediating factors (e.g., relevant knowledge and attitudes). (All of these one-page summaries are available from the first author upon request.)

All effects on behaviors or mediating factors were considered significant if: (1) they were statistically significant at the $p < .05$ level and (2) this significance was based on either the total study population or a large subgroup that was roughly at least one third of the population (e.g., males or females, one of the three major racial/ethnic groups in the U.S., or sexually experienced or inexperienced youth).

Studies sometimes reported results for multiple measures of each behavior, for different time periods, for different subpopulations, or for various combinations of the above. Thus, some studies reported one or a very small number of positive effects on behavior but also reported a large majority of results that were not significant. To avoid presenting only the positive results and to provide a more balanced overview of the results, the following rules for summarizing results were adopted.

- Regarding different measures of the same outcome behavior, all measures across all studies were rank-ordered according to their probable impact on prevalence. For example, use of condoms over 12 months was ranked higher than was condom use at first sex. Only the results from the highest ranked measure reported in each study were included in tables.
- Regarding different time periods, because very short-term effects on behaviors would have had little impact on HIV prevalence, only those results for 3 months or 6 months (depending on the behavior) or longer were included in tables.
- Regarding different subsamples, the results had to describe a subsample representing roughly one third of the sample or more (e.g., males or females) to be included.

Even these rules for summarizing results provided a more positive picture than all of the results from all studies.

However, this positive bias was at least partly offset by a different negative bias—many results presented in the studies were based on samples with insufficient power. At least half of the studies reviewed lacked statistical power to detect meaningful program effects on behavior. Moreover, the problem of insufficient power was further aggravated by the fact that studies typically had to divide their samples into various subsamples. For example, to measure the impact of the program on sexual initiation, the samples were typically restricted to those who were sexually inexperienced at baseline, and to measure impact on condom or contraceptive use, the samples were commonly restricted to those who were sexually active.

Identification of characteristics of effective programs

After analyzing the behavioral effects of the programs evaluated, we used a three-step process to identify the com-
mon characteristics of those programs that were effective at changing behavior.

First, to generate a comprehensive list of potentially important characteristics of programs, we examined reviews of health education and HIV education programs for adults, reviews of sex and HIV education programs for young adults, and individual studies of sex and HIV education programs for youth.

Second, to identify the common characteristics of curriculum content (as opposed to curriculum development and implementation), from among the 83 studies, we identified and requested 28 curricula that had the strongest evidence for positive behavioral effects; the evidence for the positive effects of the remaining 55 curricula was less strong (e.g., the results were either not significant or the evaluation designs and/or results were less strong. Of these 28 effective curricula, we successfully obtained 19.

We then coded most activities in these 19 curricula according to the risk and protective factors they addressed. Activities from different curricula addressing the same risk and protective factor were photocopied, placed in a separate binder for each factor, and subsequently reviewed. Finally, we rated each overall curriculum on each of the potentially important characteristics that was either on the list of potentially important characteristics identified in the first step or that emerged from our in-depth review. Those characteristics that were incorporated into more than 80% of the effective curricula became the eight common characteristics of effective curricula content described below.

Third, to determine more accurately the characteristics of the process for developing and implementing the effective curricula (as opposed to the content of the curricula), we coded the original research articles and reviewed any other materials (e.g., the curricula themselves) that described how the curricula were developed and implemented. These descriptions became the basis for the five characteristics describing the development of the program and the four characteristics describing the implementation of the program (described below).

Although we coded the curricula and studies as objectively as we could, our results necessarily had to reflect some degree of judgment, in part, because many studies lacked a clear and detailed description of the program development, content, and implementation.

Results

The results are divided into four sections: characteristics of the studies reviewed, impact of programs on sexual risk behaviors and pregnancy and STD rates, impact of programs on mediating factors for sexual risk behaviors, and common characteristics of effective programs.

Characteristics of the studies reviewed

Eighty-three studies met the inclusion criteria above [11–112]. Of these, 56 were conducted in the United States; 9 in other developed countries (Canada, Netherlands, Norway, Spain, and the United Kingdom); and the remaining 18 in developing countries (Belize, Brazil, Chile, Jamaica, Kenya, Mexico, Namibia, Nigeria, South Africa, Tanzania, Thailand, and Zambia).

Despite the fact that these programs were implemented throughout the world, they had numerous characteristics in common and many of them incorporated many of the characteristics of programs previously found to be associated with effectiveness [113].

For example, 52% focused on preventing only STD/HIV, 31% focused on preventing both STD/HIV and pregnancy, and 17% focused only on teen pregnancy. This greater emphasis on STD/HIV undoubtedly reflects the worldwide concern about young adults contracting HIV and the funds and other resources devoted to reducing STD/HIV transmission.

Virtually all the programs encouraged specific sexual and protective behaviors. The vast majority of the programs encouraged abstinence, but also discussed or promoted the use of condoms and/or other forms of contraception if young people chose to be sexually active. Only 7% of the programs were abstinence-only programs. All of these were in the United States. Only a few focused only on condoms. Typically, these were designed for older youth who were already likely to be sexually active.

More than four fifths of the programs (83%) identified one or more theories that formed the basis for their programs, and often specified particular psychosocial mediating factors to be changed. Social learning theory and its sequel, social cognitive theory, formed the basis for more than half (54%) of the interventions. Related theories identifying some of the same mediating factors were mentioned by substantial percentages of other studies: theory of reasoned action (19%); health belief model (12%); theory of planned behavior (10%); and the information, motivation, and behavioral skills model (10%).

Nearly all (90%) of the interventions included at least two different interactive activities designed to involve youth and help them personalize the information (e.g., role playing, simulations or individual worksheets that applied lessons to their lives).

Finally, at least 90% of the programs trained their educators before the educators implemented curriculum activities. Some of the remaining 10% of programs may also have trained their educators, but their respective reports or articles did not mention the training.

It should be noted parenthetically that some of these commonalities reflected the criteria for inclusion in this review. Only studies that addressed STD/HIV or pregnancy and were both curriculum-based and group-based were included; broad youth development programs or sex education programs that did not address STD/HIV or pregnancy at all were not included.

The one area of great variation in these programs involved length of study. While the mean length was 12 hours, the actual lengths ranged from less than one hour to 48 hours. About
two-thirds had between 2 and 15 sessions, but 11% included only 1 session, and 22% had 16 or more sessions.

To evaluate program impact, half (51%) of these studies employed an experimental design with random assignment of individual youth, classrooms of youth, or entire schools or communities, while the remaining half used a quasi-experimental design. To be included in this review, all quasi-experimental designs had to have both intervention and comparison groups and both pretest and post-test data. About 88% of all the studies used a matched-cohort design in which they linked baseline and follow-up survey data, while the others (only 12%) used unmatched pre and post cross-sectional surveys.

All measures of sexual and contraceptive behavior relied on self-reports. Although some under- and over-reporting of these behaviors undoubtedly occurred, these data are generally believed to be reasonably reliable and valid in developed countries, and often, but not always, the biases may be similar in both the intervention and control groups, especially when the data are collected confidentially by data collectors months after the end of the intervention [114]. However, this may not be the case in some developing countries, where youth are far less accustomed to talking about sexual behavior or completing questionnaires about personal behavior [115].

Measures of pregnancy and STD can be measured with laboratory tests and thereby overcome many of the problems of self-reported data. Of the 13 studies that measured impact on pregnancy, 4 used pregnancy tests; of the 10 studies that measured impact on STD, 5 used a laboratory test. The rest relied on self-reports.

As noted above, to be included in this review, studies also had to measure behavior for at least 3 months (or initiation of sex for at least 6 months). Just over half (59%) of the studies measured impact for a year or longer, while 22% measured impact for 2 years or longer. The longest study measured impact over 57 months.

In sum, as the field is maturing, increasingly large percentages of studies have used experimental designs, have used cohort designs, and have measured long-term impact on behavior. These changes as well as improvements in other areas (e.g., more rigorous and sophisticated statistical analyses) demonstrate that standards are becoming more rigorous. On the other hand, many of these studies had significant limitations such as limited explanations of the programs, problems with implementation, weak evaluation designs, measurement issues, and statistical shortcomings.

Impact of programs on sexual risk behaviors and pregnancy and STD rates

All but one of the 83 studies measured program impact on one or more of six sexual behaviors: initiation of sex, frequency of sex, number of sexual partners, condom use, contraceptive use in general, and composite measures of sexual risk-taking. A few studies reported on pregnancy and STD rates (Table 1).

Initiation of sex

An important measure of sexual activity is timing of initiation of sex. The studies reviewed demonstrate that these programs in general did not hasten the initiation of sex, and some delayed the initiation of sex. Of the 52 studies that measured impact on this behavior, 22 (42%) found that the programs significantly delayed the initiation of sex among one or more groups for at least 6 months, 29 (55%) found no significant impact, and 1 found the program hastened the initiation of sex.

Frequency of sex

A second measure of sexual activity is the frequency of sex during a specified period of time (e.g., 3 or 6 months prior to the survey). This measure includes whether or not respondents had sex at all during that period of time. Of the 31 studies that measured impact on frequency, 9 (29%) reduced the frequency, 19 (61%) found no significant change in frequency, and 3 (all in developed countries) found increased frequency among any major groups at any point in time.

Number of sexual partners

A third common measure of sexual activity is number of sexual partners during a specified period of time prior to the survey. This measure is especially important for STD transmission. Of 34 studies measuring this factor, 12 (35%) found a decrease in the number of sexual partners, while 21 (62%) found no significant impact, and 1 (3%) found a negative impact.

Condom use

Of the 54 studies measuring program impact on condom use, almost half (48%) showed increased condom use; none found decreased condom use.

Contraceptive use in general

Of the 15 studies measuring impact, 6 showed increased contraceptive use, 8 showed no impact, and 1 showed decreased contraceptive use.

Sexual risk-taking

Some studies (28) developed composite measures of sexual activity and condom use (e.g., frequency of sex without condoms or number of unprotected sexual partners). Half of them found that the programs significantly reduced sexual risk-taking. None of them found increased sexual risk-taking.

Pregnancy rates

Of the 13 studies that measured pregnancy rates, 3 found significant positive impacts, 9 found insignificant impacts, and 1 found a significant negative impact.

STD rates

Of the 10 studies that measured impact on STD rates, 2 found positive impacts, 6 found no significant impact, and 2 found negative impacts.
Overall, these studies strongly indicate that these programs were far more likely to have a positive impact on behavior than a negative impact. Across all 83 studies, two thirds (65%) had a significant positive impact on one or more of these sexual behaviors or outcomes, while only 7% had a significant negative impact on one or more of these behaviors or outcomes. Given the large number of tests of significance across all of these studies, some, but not all, of the positive and negative results undoubtedly occurred by chance. Given the large proportion of studies that found significant positive results, a few, but not all, probably occurred by chance. On the other hand, given that multiple coefficients were examined in each study, the percentage of significant negative results that was found is roughly equal to or less than the number that is likely to have occurred by chance.

One third (33%) of the programs had a positive impact on two or more behaviors or outcomes. For example, in the U.S., *Becoming a Responsible Teen* increased abstinence, reduced the number of sexual partners, increased condom use, and reduced unprotected sex [94]. Also in the U.S., the *Safer Choices* intervention delayed the initiation of sex among Hispanic youth, and increased both condom and contraceptive use among both boys and girls of all races/ethnicities [55]. In Tanzania, the *MEMA kwa Vijana* intervention both reduced the number of sexual partners among boys and increased condom use among both boys and girls.

<table>
<thead>
<tr>
<th>Outcomes measured</th>
<th>U.S. (N = 56)</th>
<th>Other developed countries (N = 9)</th>
<th>Developing countries (N = 18)</th>
<th>All countries (n = 83)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delay sex</td>
<td>(n = 30)</td>
<td>(n = 8)</td>
<td>(n = 14)</td>
<td>(n = 52)</td>
</tr>
<tr>
<td>Increased use</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1 2%</td>
</tr>
<tr>
<td>No significant results</td>
<td>15</td>
<td>6</td>
<td>8</td>
<td>29 56%</td>
</tr>
<tr>
<td>Reduced number</td>
<td>14</td>
<td>2</td>
<td>6</td>
<td>22 42%</td>
</tr>
<tr>
<td>Reduce frequency of sex</td>
<td>(n = 24)</td>
<td>(n = 2)</td>
<td>(n = 5)</td>
<td>(n = 31)</td>
</tr>
<tr>
<td>Increased frequency</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>3 10%</td>
</tr>
<tr>
<td>No significant results</td>
<td>15</td>
<td>1</td>
<td>3</td>
<td>19 61%</td>
</tr>
<tr>
<td>Reduced frequency</td>
<td>7</td>
<td>0</td>
<td>2</td>
<td>9 29%</td>
</tr>
<tr>
<td>Reduce number of partners</td>
<td>(n = 26)</td>
<td>(n = 0)</td>
<td>(n = 8)</td>
<td>(n = 34)</td>
</tr>
<tr>
<td>Increased numbers</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1 3%</td>
</tr>
<tr>
<td>No significant results</td>
<td>16</td>
<td>0</td>
<td>5</td>
<td>21 62%</td>
</tr>
<tr>
<td>Reduced number</td>
<td>9</td>
<td>0</td>
<td>3</td>
<td>12 35%</td>
</tr>
<tr>
<td>Increase condom use</td>
<td>(n = 37)</td>
<td>(n = 5)</td>
<td>(n = 12)</td>
<td>(n = 54)</td>
</tr>
<tr>
<td>Reduced use</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0 0%</td>
</tr>
<tr>
<td>No significant results</td>
<td>19</td>
<td>4</td>
<td>5</td>
<td>28 52%</td>
</tr>
<tr>
<td>Increased use</td>
<td>18</td>
<td>1</td>
<td>7</td>
<td>26 48%</td>
</tr>
<tr>
<td>Increase contraceptive use</td>
<td>(n = 11)</td>
<td>(n = 2)</td>
<td>(n = 2)</td>
<td>(n = 15)</td>
</tr>
<tr>
<td>Reduced use</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1 7%</td>
</tr>
<tr>
<td>No significant results</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>8 53%</td>
</tr>
<tr>
<td>Increased use</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>6 40%</td>
</tr>
<tr>
<td>Reduce sexual risk taking</td>
<td>(n = 25)</td>
<td>(n = 1)</td>
<td>(n = 2)</td>
<td>(n = 28)</td>
</tr>
<tr>
<td>Increased risk</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0 0%</td>
</tr>
<tr>
<td>No significant results</td>
<td>11</td>
<td>1</td>
<td>2</td>
<td>14 50%</td>
</tr>
<tr>
<td>Reduced risk</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>14 50%</td>
</tr>
<tr>
<td>Reduce pregnancy: self-report</td>
<td>(n = 6)</td>
<td>(n = 3)</td>
<td>(n = 0)</td>
<td>(n = 9)</td>
</tr>
<tr>
<td>Increased number</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1 11%</td>
</tr>
<tr>
<td>No significant results</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>6 67%</td>
</tr>
<tr>
<td>Reduced number</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2 22%</td>
</tr>
<tr>
<td>Reduce pregnancy: laboratory test</td>
<td>(n = 2)</td>
<td>(n = 0)</td>
<td>(n = 2)</td>
<td>(n = 4)</td>
</tr>
<tr>
<td>Increased number</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0 0%</td>
</tr>
<tr>
<td>No significant results</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>3 75%</td>
</tr>
<tr>
<td>Reduced number</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1 25%</td>
</tr>
<tr>
<td>Reduce STDs: self-report</td>
<td>(n = 3)</td>
<td>(n = 0)</td>
<td>(n = 2)</td>
<td>(n = 5)</td>
</tr>
<tr>
<td>Increased number</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1 20%</td>
</tr>
<tr>
<td>No significant results</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>4 80%</td>
</tr>
<tr>
<td>Reduced number</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0 0%</td>
</tr>
<tr>
<td>Reduce STDs: laboratory test</td>
<td>(n = 4)</td>
<td>(n = 0)</td>
<td>(n = 1)</td>
<td>(n = 5)</td>
</tr>
<tr>
<td>Increased number</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1 20%</td>
</tr>
<tr>
<td>No significant results</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2 40%</td>
</tr>
<tr>
<td>Reduced number</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2 40%</td>
</tr>
</tbody>
</table>

* None of the studies measured all of the outcomes. Thus, the sum of the number of negative, nonsignificant and positive results for each outcome is less than the total number of studies.
These effects in these three studies are particularly noteworthy, because all three studies employed experimental designs and measured impact on behavior for at least 1 year. More generally, studies indicate that it is possible both to reduce sexual behavior and to increase condom or contraceptive use with the same program.

While the positive effects of some curriculum-based programs lasted only a few months, the effects of other programs lasted for years. For example, the MEMA kwa Vijana [82] intervention found positive behavioral effects over a 36-month period and Safer Choices [26] found positive behavioral effects over a 31-month period.

These findings were remarkably robust. The patterns of findings remained the same regardless of the type of experimental design used. Whereas 63% of the 41 studies with a quasi-experimental design had a significant positive impact on one or more behaviors, 66% of the 42 studies with an experimental design had a positive impact on one or more behaviors.

The patterns of findings were also similar in both developed and developing countries. Programs were effective with both low- and middle-income youth, in both rural and urban areas, with girls and boys, with different age groups, in school, clinic, and community settings (results not shown). (This does not mean that the same program was effective with all of these groups in all of these areas, but simply that different programs were effective with all of these groups in all of these areas.)

And finally, while there were only six studies that focused only on abstinence (all in the U.S.), there were a few positive results (and one negative result) for these programs, just as there were many positive results and a few scattered negatives results for the far more numerous programs that commonly emphasized both abstinence and condom or contraceptive use.

Replication of studies

A critically important question is whether or not a program that has been found to be effective when designed, implemented, and evaluated by a well-funded and highly skilled research team, will subsequently be effective when implemented by others in other communities. Four interventions in the United States have now been evaluated between two and five times [18,47,48,50–52,54,94,95,97,98,111,112,116,117]. These replications reveal that these curricula consistently had similar positive behavioral effects when they were replicated, provided (1) all activities were implemented as designed and (2) they were implemented in the same type of setting and with similar populations of youth. When many activities were omitted or the setting was changed (e.g., from voluntary Saturday programs to required in-school programs), the curricula were less likely to have a positive effect.

Impact of programs on mediating factors for sexual risk behaviors

Although the review of the studies above provides strong evidence that some programs had an impact on sexual risk behaviors, without the results of the mediating factors, it does not specify how or why these programs had an impact. Those questions can be partially answered by examining programmatic impact on the mediating factors that programs attempted to change in order to change behavior. Table 2 provides the results for 71 different mediating factors that were measured by one or more of the 83 studies. Those factors that are lightly shaded meet two criteria and therefore have stronger evidence that programs can modify them: (1) at least three programs significantly improved them and (2) at least half of the studies that measured them found significant improvements.

Of those programs that measured impact, most increased knowledge about HIV, STDs, and pregnancy (including methods of preventing STD/HIV and pregnancy). Half of the 16 studies that measured impact on perceived HIV risk were effective at increasing this perceived risk. More than 60% of the many studies measuring impact on values and attitudes regarding any sexual topic were effective in improving the measured values and attitudes. About 40% of the 29 studies that measured impact on perceived peer sexual behavior and norms significantly improved these perceptions. More than half of those studies that measured impact on self-efficacy to refuse unwanted sex improved that self-efficacy, and more than two thirds increased self-efficacy to use condoms. Ten of 16 programs increased motivation or intention to abstain from sex or restrict the number of sex partners, and 10 of 14 programs increased intention to use a condom. Eight of 11 programs increased communication with parents or other adults about sex, condoms, or contraception. In contrast, less than 30% of the programs had a positive impact on use of drugs or alcohol, perhaps in part, because reducing use of alcohol or drugs was not the focus of these programs.

In sum, the evidence was strong that many programs had positive effects on relevant knowledge, awareness of risk, values and attitudes, self-efficacy, and intentions—the very factors specified by many psychosocial theories as being the determinants of behavior. Furthermore, all of these factors have been demonstrated empirically to be related to their respective sexual behaviors [117]. Thus, it appears highly likely that changes in these factors contributed to the changes in sexual risk-taking behaviors.

Characteristics of effective curricula

The in-depth analyses of effective programs identified 17 characteristics that described these programs and are presented in Figure 1. Three types of evidence suggest that these characteristics may have contributed to the success of these programs. First, a large majority of the effective programs shared most of these characteristics. Second, programs that incorporated these characteristics were more likely to change behavior positively than programs that did not incorporate many of these characteristics. Third, several studies involved a comparison of the impact of skill-based curricula that incorporated all (or nearly all) of these characteristics with the impact of knowledge-based curricula that did not incorporate many of these
Characteristics. Consistently, the skill-based programs were more effective at changing behavior than were the knowledge-based programs [26,95].

Although nearly all of the effective curricula incorporated nearly all of the effective characteristics, and although curricula with nearly all of the characteristics were highly likely to be effective, having most of the 17 characteristics present in a curriculum did not ensure significant changes in reported behavior. In addition, a few curricula that did not appear to

Table 2
Number of programs having effects on mediating factors that may affect sexual behavior or condom or contraceptive use

<table>
<thead>
<tr>
<th>Pos NS Neg</th>
<th>Intention to discuss condoms with partner</th>
<th>Intention to abstain from sex, or restrict sex or partners</th>
<th>Intention to use a condom</th>
<th>Intention to avoid unprotected sex (including perceived likelihood of having sex)</th>
<th>Communication with partner re AIDS, STDs, and past partners</th>
<th>Communication with boys/girlfriends or partners regarding abstinence or condoms use</th>
<th>Communication with parents or other adults about sex, condoms, or contraception</th>
<th>Comfort talking with parents about sex, condoms, or contraception</th>
<th>Other possible mediating behaviors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>10</td>
<td>6</td>
<td>0</td>
<td>10</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Knowledge</td>
<td>Overall knowledge of sexual issues</td>
<td>7</td>
<td>2</td>
<td>0</td>
<td>Knowledge of pregnancy</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>Knowledge of STD</td>
</tr>
<tr>
<td>Perceived risk</td>
<td>Perception of pregnancy risk</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>Perception of STD risk</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>Perception of HIV risk</td>
</tr>
<tr>
<td>Personal values and attitudes</td>
<td>Values about sex/abstinence</td>
<td>14</td>
<td>10</td>
<td>0</td>
<td>Regret about initiating sex</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>Attitude about pressuring someone to have sex (including right to say no to sex)</td>
</tr>
<tr>
<td>Perceived peer values and behavior</td>
<td>Attitude toward risky sexual behavior and AIDS prevention</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>Self-appraisal to use condoms</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>Attitudes toward HIV+ people (including interacting with them)</td>
</tr>
<tr>
<td>Perceived partner values</td>
<td>Perceived partner norms and reaction to condom use</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>Self-efficacy and skills</td>
<td>Self-efficacy to show love and affection without sex</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Motivation/Intentions</td>
<td>Condom use skills</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>Self-efficacy to avoid STD/HIV risk and risk behaviors (e.g., to abstain or use condoms)</td>
<td>7</td>
<td>9</td>
<td>0</td>
<td>General sexual negotiation skills</td>
</tr>
<tr>
<td>Intention to discuss AIDS, STDs, and past partners with new partner</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a None of the studies measured all of the outcomes. Thus, the sum of the number of negative, nonsignificant and positive results for teach outcome is less than the total number of studies.

b Pos = Positive (desirable) effect on factor; NS = Nonsignificant; Neg = Negative (undesirable) effect on factor.

c Lightly shaded factors meet two criteria and therefore have stronger evidence that programs can modify them: (1) at least three programs significantly improved them and (2) at least half of the studies that measured them found significant improvements.
incorporate all 17 characteristics nevertheless had a positive impact.

It should also be recognized that these characteristics of effective programs are not the only factors that determine whether or not programs will have an impact on behavior. Other factors, such as the saliency of unintended pregnancy, HIV or other STDs and the existing knowledge, values, attitudes, and skills of young also have an impact. Thus, for example, if HIV is a very salient issue in a community and youth in that community lack basic information about how HIV is transmitted, the chances of infection during unprotected sex, and methods of protection, then programs that provide this basic information may have an impact on behavior, even if they do not incorporate all of the 17 characteristics.

The 17 characteristics can logically be divided into three categories, namely those describing: (1) the development of the curricula, (2) the overall design and teaching strategies of the curricula themselves, and (3) the implementation of the curricula. As noted above, they are presented in Figure 1.

These characteristics can be used to assess and select curricula [118], to adapt or improve them and even to develop them from scratch. They can also be used to guide implementation.

The Process of Developing the Curriculum

1. Involved multiple people with different backgrounds in theory, research and sex/HIV education to develop the curriculum
2. Assessed relevant needs and assets of target group
3. Used a logic model approach to develop the curriculum that specified the health goals, the behaviors affecting those health goals, the risk and protective factors affecting those behaviors, and the activities addressing those risk and protective factors
4. Designed activities consistent with community values and available resources (e.g., staff time, staff skills, facility space, and supplies)
5. Pilot-tested the program

The Contents of the Curriculum Itself

Curriculum Goals and Objectives
1. Focused on clear health goals – the prevention of STD/HIV and/or pregnancy
2. Focused narrowly on specific behaviors leading to these health goals (e.g., abstaining from sex or using condoms or other contraceptives), gave clear messages about these behaviors, and addressed situations that might lead to them and how to avoid them
3. Addressed multiple sexual psychosocial risk and protective factors affecting sexual behaviors (e.g., knowledge, perceived risks, values, attitudes, perceived norms, and self-efficacy)

Activities and Teaching Methodologies
4. Created a safe social environment for youth to participate
5. Included multiple activities to change each of the targeted risk and protective factors
6. Employed instructionally sound teaching methods that actively involved the participants, that helped participants personalize the information, and that were designed to change each group of risk and protective factors
7. Employed activities, instructional methods and behavioral messages that were appropriate to the youths’ culture, developmental age, and sexual experience
8. Covered topics in a logical sequence

The Implementation of the Curriculum

1. Secured at least minimal support from appropriate authorities such as ministries of health, school districts or community organizations
2. Selected educators with desired characteristics (whenever possible), trained them and provided monitoring, supervision and support
3. If needed, implemented activities to recruit and retain youth and overcome barriers to their involvement, e.g., publicized the program, offered food, or obtained consent
4. Implemented virtually all activities with reasonable fidelity

Figure 1. Characteristics of effective curriculum-based programs.

Conclusions and Recommendations

Many of these studies had significant limitations. For example, few described their respective programs adequately; none studied programs for youth engaging in same-sex behavior; some had problems with implementation; a few had relatively weak quasi-experimental designs; an unknown number had measurement problems; many were statistically underpowered; most did not adjust for multiple tests of significance; few measured impact on either STD or pregnancy rates; and still fewer measured impact on STD or pregnancy rates with biomarkers. And, of course, there are inherent publication biases that affect the publication of studies—researchers are more likely to try to publish articles if positive results support their theories and programs and journals are more likely to accept articles for publication if results are positive. Fortunately, some of these biases counteract each other.

Despite these limitations, the evidence for the positive impact on behavior of curriculum- and group-based sex and HIV education programs for adolescents and young adults is quite strong and encouraging. Two thirds of the programs had a significant positive impact on behavior. Many either delayed or reduced sexual activity or increased condom or contraceptive use or both. At least 10 interventions had long-term behavioral
effects lasting 2 or more years; some lasted for close to 3 or more years—as long as the effects were measured. Most programs also increased psychosocial mediating factors that are known to be related to sexual behavior. These studies help explain how these programs are effective.

The evidence is also strong that these programs in general did not have negative effects. In particular, they did not increase sexual behavior, as some people have feared they might. Of the 52 studies that measured impact on initiation of sex, only one significantly hastened the initiation of sex. Given the large number of studies and tests of significance for that outcome, that could have occurred by chance. The few other scattered negative findings among both abstinence-only and comprehensive sex and HIV education programs may also have occurred by chance. The evidence is dramatically stronger that these programs had positive effects on sexual behavior.

The effects of these programs were quite robust. They were just as likely, if not more likely, to be effective in developing countries as they were to be effective in the U.S. or other developed countries. They were effective in urban and rural areas, low- and middle-income communities, and school and community settings, with advantaged and disadvantaged youth, males and females, different racial and ethnic groups, younger and older youth, and sexually experienced and inexperienced youth. There is some indication that they were especially effective with youth who were most likely to engage in unprotected sex with multiple partners and thus were at highest risk of HIV, other STDs and pregnancy. Of course, the exact same program was not implemented with all of these groups; rather programs were appropriately designed or tailored for some of these groups.

Robustness was also demonstrated by replication studies. When three programs were replicated with fidelity in different locations throughout the United States, but in the same type of setting, the original positive effects were confirmed. This is very encouraging and suggests that effective programs can remain effective when they are implemented by people in other communities.

Given that many programs reduced sexual behavior and/or also increased condom or contraceptive use, they logically would reduce both sexually transmitted disease and pregnancy. The results of the few studies that measured impact on STD or pregnancy, however, did not produce many significant positive effects. The lack of consistent positive effects may have been caused, in part, by sample sizes that were too small, by other methodological limitations, by significant changes in behavior that were too small or too short term to produce marked changes in STD or pregnancy, or possible failure to change those behaviors that have the strongest impact on STD or pregnancy rates.

Thus, while these programs alone cannot solve the problems of STD, HIV, and unintended pregnancy, many of them can change sexual and protective behaviors in desired directions and they can be an important component in larger more comprehensive initiatives.

These conclusions support several programmatic and research conclusions:

- Communities should implement curriculum-based sex and HIV education programs, preferably those proven to be effective with similar populations or those incorporating the 17 characteristics of effective programs.
- Because these programs can reduce sexual risk-taking by a modest amount, communities should not rely solely on these programs to address problems of HIV, other STDs, and pregnancy, but should view them as an important component in a larger initiative that can reduce sexual risk-taking behavior to some degree.
- More rigorous studies of promising programs need to be conducted, especially in developing countries and with groups at highest risk, because there are gaps in these areas in the existing literature.
- Evaluations can and should use randomized experimental designs. They have been used very successfully even in the poorest regions of the world.
- Sample sizes should be sufficiently large to have adequate statistical power for important statistical analyses, including those among subgroups. Conducting studies that are substantially underpowered is unfair to their respective programs and may mislead the field.
- Laboratory tests rather than self-reported data have been used for measuring pregnancy and STD rates, and should be used whenever feasible. These studies are particularly important given that many programs have changed behavior, but few have changed pregnancy and STD rates. This apparent inconsistency needs to be further understood.
- Researchers should determine more rigorously which mediating factors have the greatest impact on behavior in different cultures and which educational strategies and activities are most effective at changing these factors both across and within cultures. More generally, studies should try to assess which of the 17 characteristics are most important and what kinds of adaptations can be safely made without jeopardizing effectiveness.
- Researchers should always try to publish negative results, provided that the studies are well done, so that the literature does not become biased.
- Published results of evaluations should provide more complete descriptions of their programs so that reviews can better assess their characteristics and understand why some programs were effective and other were not.
- Formal meta-analyses of all of these studies should be conducted so that they can overcome some of the limitations of the individual studies (e.g., insufficient statistical power).

If these recommendations are implemented, the success of these programs and our knowledge of these programs will continue to progress, as they have for the last 25 years.
References


Ross D. MEMA Kwa Vijana: Randomized controlled trial of an adolescent sexual health programme in rural Mwanza, Tanzania. London: London School of Hygiene and Tropical Medicine, 2003.


