

The Impact of the Postponing Sexual Involvement Curriculum Among Youths in California

By Douglas Kirby, Meg Korpi, Richard P. Barth and Helen H. Cagampang

Postponing Sexual Involvement (PSI) is a widely implemented middle school curriculum designed to delay the onset of sexual intercourse. In an evaluation of its effectiveness among seventh and eighth graders in California, 10,600 youths from schools and community-based organizations statewide were recruited and participated in randomly assigned intervention or control groups; the curriculum was implemented by either adult or youth leaders. Survey data were collected before the program was implemented, and at three months and 17 months afterward. At three months, small but statistically significant changes were found in fewer than half of the measured attitudes, behaviors and intentions related to sexual activity; at 17 months, none of these significant positive effects of the PSI program had been sustained. At neither follow-up were there significant positive changes in sexual behavior: Youths in treatment and control groups were equally likely to have become sexually active, and youths in treatment groups were not less likely than youths in control groups to report a pregnancy or a sexually transmitted infection. The evaluation suggests that PSI may be too modest in length and scope to have an impact on youths' sexual behavior.

(Family Planning Perspectives, 29:100–108, 1997).

Concern about high rates of unintended pregnancy and sexually transmitted disease (STD) infection among adolescents has led to the implementation in many middle and high schools of sexuality education programs designed to delay the onset of sexual intercourse. One such curriculum, Postponing Sexual Involvement (PSI),¹ aims to support adolescents in delaying sexual activity by helping them understand the various social pressures that encourage adolescent sexual activity and by teaching teenagers skills that will enable them to set limits, resist peer pressure, be assertive in saying “no” to sex and develop nonsexual ways to express their feelings.

The PSI program is probably the most widely implemented middle school cur-

riculum of its kind. It is brief and takes up little class time and, given its focus on postponing sex, has broad appeal to parents and schools. Moreover, an Atlanta-based evaluation of the PSI program implemented in combination with a five-session human sexuality course suggested that the curriculum delayed the initiation of first intercourse.²

However, serious questions have been raised about the quality of the intervention. For example, is the curriculum too modest in length, does it include enough practice in the skills that it attempts to teach, and do the slides or videos that it employs have appeal for young people?

Furthermore, two important methodological limitations of the Atlanta evaluation have been noted.³ First, students were not randomly assigned to treatment and control groups: Youths living in one geographic area received the intervention, and they were compared with youths living in other geographic areas who did not receive the program. Thus, while analytic procedures controlled for some background characteristics and the two groups appeared to have been well matched,

other uncontrolled factors may have differentially affected the two samples.

Second, the results of the Atlanta evaluation were biased slightly, because a small number of youths in the treatment group who initiated intercourse during the semester in which they participated in the program were excluded from the statistical analysis, while no comparable youths from the comparison group were omitted.

See also pp. 109–114.

The California Replication

From 1992 to 1994, in an effort to reduce teenage pregnancy statewide, the California Office of Family Planning funded the Education Now and Babies Later (ENABL) initiative.⁴ Composed of 28 projects coordinated by an array of local non-profit, educational, health and social service agencies, ENABL represented the largest statewide pregnancy prevention effort ever initiated.

The ENABL initiative included the PSI curriculum and school- and community-wide activities (such as flyer distributions, assemblies, rallies and fairs) designed to promote healthy alternatives to sexual activity, involve large numbers of youths in the ENABL campaign and increase acceptance of the program and of its messages. The initiative also included a statewide media campaign and provided youths with referral information for health and other social services.

Collectively, the ENABL projects delivered the PSI curriculum to approximately 187,000 youths in schools and community settings in 31 California counties. Ninety percent of the PSI programs were taught by adults (mostly professional educators or, occasionally, college interns). Ten percent of the programs were taught by youth leaders who were teenagers slightly older than those participating in

Douglas Kirby is director of research and Meg Korpi is senior research scientist at ETR Associates, Santa Cruz, Calif.; Richard P. Barth is Hutto Patterson Professor and Helen H. Cagampang is research specialist at the School of Social Welfare, University of California, Berkeley. This research was supported principally by the California Office of Family Planning, with additional support from the Cowell Foundation, the Packard Foundation and the Stuart Foundations.

the evaluation and who were trained to lead the intervention groups. Youth leaders were always accompanied by adult observers when presenting the program.

Seventeen ENABL projects also utilized PSI for Parents, a companion to the youth curriculum. It was designed to help parents reinforce their children's learning experiences regarding postponing sexual involvement. As part of ENABL, PSI for Parents was generally given in one 90–120-minute session. However, only about 5% of the parents of youths in this study received PSI for Parents.

The California evaluation was designed to measure PSI's impact on the occurrence of first intercourse and to examine the beliefs, attitudes and intentions that might mediate the initiation of sexual intercourse. The evaluation tested the effectiveness of implementing the program in both school and community settings, and examined school-based, adult-led interventions as well as those led by teenagers. In some agencies, the evaluation also assessed the impact of schoolwide and community-wide ENABL activities. The evaluation also examined the program's differential impact on sexual behavior according to participant's gender, grade, racial and ethnic background, and prior sexual experience.

Methods

Contractors

School districts, health departments and community-based organizations applied to the California Office of Family Planning to obtain funding to implement PSI and ENABL. Twenty-eight organizations were given contracts to use trained intervention leaders to implement the PSI program in school and community settings. These organizations were selected because they provided service to communities with high teenage birthrates, as well as for geographic and ethnic diversity and the ability to deliver a program like PSI.

Accordingly, the sample in this study is diverse, but it is likely to be more representative of youths in areas with higher rates of sexual risk behavior and higher teenage birthrates than of all California youths. Twenty-one of the 28 selected organizations completed all of the requirements of the evaluation and are included in the results. In all, 56 middle or junior high schools and 17 community-based agencies participated in the evaluation.

The PSI Program

The PSI curriculum consists of five sessions, 45–60 minutes in length, delivered in classroom or small group settings. Ses-

sion I focuses on the risks of early sexual involvement and helps youths explore the reasons that teenagers have sex and the reasons why they might choose to wait. Session II helps young people understand and resist the social pressures that can lead to early sexual involvement. Session III identifies peer pressures that can affect teenagers' sexual behavior and helps teenagers determine their own limits for physically expressing affection. Session IV teaches assertive responses to help teenagers resist pressure to engage in sex. Session V provides reinforcement of the material learned in previous sessions. The PSI intervention included class discussions, group activities, use of videos or slides and a small amount of role playing.

PSI was implemented in addition to whatever standard sexuality curriculum an individual school offered. Thus, students in both treatment and control groups were likely to receive some instruction in sexuality. However, the vast majority of students in the control groups were not offered an additional, specialized sexuality curriculum comparable to PSI. Instead of PSI, these students typically received instruction in some other topic area.

Research Design

Three research designs, representing three levels of random assignment, were used to evaluate the effectiveness of PSI and the ENABL program. Each design included some level of random assignment, and the collection of survey data before the delivery of the intervention and again 17 months later.

In the first design, students within selected schools were randomly assigned by classroom to either a youth-led intervention, an adult-led intervention or a no-intervention control. This design also involved a few contractors who implemented the adult-led PSI program but did not offer the youth-led intervention. Thus, students in this group were randomly assigned by classroom to only two conditions, either the adult-led intervention or a control group. In this classroom-randomization design, survey data were collected from students three months after baseline, as well as at baseline and at the 17 month follow-up.

In the second design, entire schools were randomly assigned to either intervention or control conditions; intervention schools received adult-led PSI as well as various schoolwide activities in support of the ENABL initiative, while control schools received the standard sexuality education curricula; data were collected at baseline and 17 months after baseline only.

In the third design, youths were recruited from community-based agencies and were randomly assigned individually to either an adult-led intervention group or a no-intervention control; data were collected at baseline and 17 months after baseline only.

PSI stands on its own as a sexuality education curriculum. However, in the Atlanta implementation, it was preceded by a five-session course in human sexuality and decision-making.⁵ In order to make the California intervention similar to the one in Atlanta, youths in California were required to receive instruction in human sexuality before participating in PSI. However, the specific curriculum used in Atlanta was not available to the public. Therefore, a different curriculum, covering similar subject matter, was used in the California replication.

We made an intensive effort to eliminate schools in which students in the control group might have received the PSI intervention. It is possible, for example, that an adolescent assigned to a control group at one of the community agencies might have received the intervention at school. However, the number of such students would be too small to influence the overall results of the study.

Sample

A total of 10,600 youths received parental consent to participate in this study; 75% completed both the baseline and the 17-month follow-up surveys. Among youths in the first research design, 4,234 (91%) also completed a three-month posttest. After surveys with incomplete or inconsistent data were eliminated,* the final sample included 7,340 youths who completed the baseline and follow-up surveys, 3,834 of whom had also completed the three-month posttest survey. Survey completion rates were similar for youths in both the intervention and control groups.

Among youths who were lost from the original sample for any reason, there were no significant differences between those lost from the intervention and those lost from the control group (about 1% more were lost from the intervention group than

*An extensive multistage process was used to clean the data. Surveys were eliminated if there was evidence of a "set" response pattern (e.g., the participant always selected "3" on a scale of 1 to 5), or if incompatible responses were evident (e.g., responses regarding having had sexual intercourse were inconsistent). (See: D. Kirby et al., *Evaluation of Education Now and Babies Later (ENABL): Final Report*, Family Welfare Research Group, Berkeley, Calif., 1995.) Six percent of cases were removed during data cleaning, leaving 69% of the original 10,600 cases in the final 17-month sample.

from the control group).^{*} We report findings for the full three-month and 17-month samples here. However, we also compared our results for these full samples with those from the smaller sample for which we had three waves of data, and found that they were consistent.

Measures

We drew upon previous research in the field for our outcome measures. On occasion, if there were no appropriate scales available, we developed our own items. Measures were reviewed by several professionals in the field and extensively pilot-tested with students who completed the draft questionnaire and then participated in focus groups to further discuss and refine the survey items. The questionnaire was also translated into Spanish.

The main outcome of interest was whether a teenager had become sexually active subsequent to the intervention. We also asked adolescents whether they had tried to initiate sex or persuade someone to engage in intercourse, as well as whether they had been the recipient of such pressures. Among sexually active youths, we examined frequency of intercourse, number of sexual partners and use of contraceptives to address beliefs that PSI had the potential to affect these behaviors as well. Finally, we measured pregnancy rates and rates of reported sexually transmitted diseases.

We conducted factor analyses of all three waves of survey data to more fully understand the underlying structure of the mediating variables. These analyses resulted in the creation of seven multi-item scales. Items that did not clearly fit into any of the scales were treated separately in later analyses. We used these single items and the seven scales to measure a range of variables that are thought to mediate adolescent decision-making regarding sexual behavior.

• *Beliefs about sexual activity.* A six-item scale measured respondents' beliefs about how they and their peers viewed the timing and circumstances of first intercourse. Participants were asked to respond on a contin-

uum of agreement to six statements such as "My best friends think that people my age should wait until they are older to have sex" and "Most students at my school think it's OK for people my age to have sex with a serious boyfriend or girlfriend." An additional four-item scale addressed beliefs about sexual pressure. Respondents were also asked to estimate the percentage of girls and boys in their school who had had sex; two additional single items addressed other beliefs about the inevitability of teenage sexual activity and whether it is possible to say no to sex without hurting the feelings of the other person.

• *Reasons to have sex or abstain.* An eight-item scale assessed possible reasons adolescents might have to postpone sex (e.g., "I would not have sex now because I'm waiting for the right person"), and a six-item scale measured possible reasons for initiating sex (e.g., "I would have sex now to feel accepted and loved").

• *Beliefs about sex and the media.* Three single items measured teenagers' beliefs regarding the extent and impact of media images about sex.

• *Parental communication.* Three single items measured whether respondents had spoken over the past year with a parent or guardian about sex.

• *Self-efficacy.* Self-efficacy in declining sex was measured with a four-item scale in which respondents were asked to indicate the degree to which they felt certain they could refuse sex in different situations. (For example, they were asked: "You are alone with a boy or a girl. You start to kiss and touch and it is hard to stop. How sure are you that you could keep from having sex?") An additional item addressed respondents' ability to express affection in a nonsexual way.

• *Behavioral intentions.* A four-item scale measured teenagers' intentions to engage in sexual activity in the future (e.g., "When it comes to sex, I have already decided 'how far' I will go"). In addition, at the 17-month follow-up, youths who had never had sexual intercourse were asked if they intended to wait until they were older to have sex, and those who had had intercourse were asked if they intended to wait before they have sex again.

All survey items examining beliefs, attitudes and intentions were recoded so that a higher score corresponded with a more desirable outcome (more conducive to postponing sexual involvement). Cronbach's alpha was used for each wave of data to calculate the interitem reliability of each scale. Six of the seven scales had coefficients that exceeded .70 at all three

survey points, while the seventh scale (beliefs about sexual pressure) had coefficients exceeding .70 for two of the three time periods. Across all scales and time periods, the mean alpha coefficient was .82, indicating acceptable internal consistency of the measures.

Analytic Procedures

To control for chance differences between groups at baseline, we calculated change scores in the outcome variables over time (posttest score minus pretest score) and compared the scores of the treatment and control groups using t-tests. This procedure eliminated the need to use analysis of covariance to control for other differences at baseline.⁶ If calculation of change scores was not feasible (e.g., when examining the impact of PSI on the frequency of sexual intercourse among youths who initiated intercourse after the pretest), the posttest scores of treatment-group participants were compared with those of the control group. Chi square tests were used for categorical data, and t-tests were used for continuous data.

We set the level of statistical significance at $p < .01$ because of our relatively large sample size and because of the large number of statistical tests that we conducted. When we examined our data at a less conservative level, those findings significant at $p < .05$ but not significant at $p < .01$ were often in inconsistent directions, suggesting they were chance occurrences.

Characteristics of the Sample

The mean age of the youths in the sample was 12.8 years and the mean grade level was 75 (Table 1). Males represented 42–45% of the participants. The sample was ethnically diverse, and race and ethnicity varied across settings: Among youths receiving adult-led PSI in any setting, 27–32% were Hispanic, approximately 40% were white, 9% were black and 12–14% were Asian or Pacific Islander, while participants who received youth-led PSI were more likely to be Hispanic (46–49%) and less likely to be white (21%). Teenagers recruited from community-based agencies were most likely to be Asian or Pacific Islander (47–52%), and least likely to be black (2–3%); approximately 20% were Hispanic, and 5–10% were white (not shown). Across all settings, almost 90% of youths lived with their mother or stepmother, while almost two-thirds had a male parental figure in the home.

Some 35–39% of youths reported ever having had a serious romantic involvement. No more than 11% of youths had

^{*}We compared youths who were lost from the intervention group with those who were lost from the control group on a variety of pretest characteristics; the two groups were very similar. Using two-way analysis of variance, we examined the interaction effects between treatment group (intervention vs. control) and attrition (whether or not the 17-month follow-up survey had been completed) in relation to background characteristics, mediating variables and outcome variables at pretest. These analyses revealed no significant interaction effects at the $p < .01$ level for age, gender, family background characteristics, risk factors such as alcohol use or low school grades, or mediating or outcome variables.

ever had sex. On average, sexually experienced youths had had sex only about 2–3 times during the preceding year. Less than 1% had ever been pregnant or caused a pregnancy, and a comparable proportion had had an STD.

Nearly half of all youths reported having made a decision to place limits on their sexual activity; only 3–6% of all youths indicated that they had tried to persuade someone to have sex with them in the last three months, but 10–17% reported having been the target of such efforts (not shown).

There were relatively few statistically significant differences at baseline between treatment and control groups across the various randomization schemes; when significant differences did occur, they were very small. All statistically significant differences between treatment and control groups occurred among youths receiving adult-led PSI and occurred in the design in which entire schools were randomly assigned. Youths in the intervention group from the schoolwide randomization were more likely than control youths to be Hispanic; they received slightly higher grades in school, were less likely to speak only English in the home and had mothers with less education. At baseline, these youths were also more likely to have ever had sex, and those who were sexually active had had slightly more sexual partners (not shown). We statistically controlled for the pretest differences in sexual behavior. Nonetheless, results from this research design should be interpreted with some caution.

Results

We examined differences between youths in all treatment and control groups across all research designs, for all mediating variables and for all variables measuring sexual activity. Findings for all variables are displayed separately in the accompanying tables for youth-led classroom and adult-led classroom PSI, schoolwide ENABL and community-based PSI. We present these findings because we feel it is important to document the consistency of our results across settings. However, because three-month data were collected only for the classroom research design, and to keep the presentation of results as straightforward as possible, we describe in the text, unless otherwise noted, only the findings from the classroom research design.

Mediating Variables

• *Beliefs about sexual activity.* At the three-month posttest, teenagers in the youth-led intervention but not those in the adult-led

intervention were significantly more likely than their counterparts in the control group to believe that they and their peers endorsed postponing sex (Table 2, p. 104). The difference in the change scores between the youth-led treatment and control groups was 0.08, an effect size of 0.15.* This difference, however, was not apparent at the 17-month follow-up.

At the three-month posttest, adolescents in both the youth-led and the adult-led classroom intervention were significantly less likely than their control group counterparts to believe that becoming sexually active during the teenage years was inevitable; these differences, however, were not significant at 17 months. Compared with their pretest responses, teenagers participating in the classroom intervention disagreed more at posttest with the statement “most teens are going to have sex, no matter what,” while youths in the control groups agreed more with this statement at posttest than at pretest. The differences in change scores between the treatment and control groups for the adult-led and youth-led interventions were 0.14 and 0.12, corresponding to effect sizes of 0.15 and 0.13, respectively. Findings among students who had not had sex at pretest were similar to the results reported above for all youths.

There were no statistically significant differences between treatment and control groups at either the three- or 17-month posttest in teenager’s beliefs about sexual pressure, in their estimates of the proportion of their peers who are sexually active or in the belief that it is possible to decline sex without hurting the other person’s feelings.

• *Reasons to have sex or abstain.* At the three-month posttest, youths in the adult-led intervention checked significantly more reasons to refrain from sex than did those in the corresponding control group (a difference of 0.06), but the effect size was small (0.15). This difference was no longer

Table 1. Means and percentages for background characteristics of students participating in treatment and control groups, by type of intervention, Postponing Sexual Involvement evaluation, California, 1992–1994

Characteristic	Youth-led		Adult-led	
	Treatment	Control	Treatment	Control
Mean age	12.8	12.9	12.8	12.8
Mean grade level	7.5	7.5	7.5	7.5
% male	44.6	42.7	41.7	42.4
Race/ethnicity (%)†				
American Indian	4.0	3.6	5.3	5.5
Asian/Pacific Islander	8.4	7.5	11.7	13.7
Black	9.8	10.1	8.9	8.7
Hispanic	46.4	48.9	31.9	27.4*
White	21.2	20.7	38.5	39.6
Other	8.1	6.9	7.2	8.0
% living with mother	87.2	87.3	88.8	88.5
% living with father	60.9	62.9	63.9	63.6
Mean academic grades‡	1.7	1.7	1.6	1.5*
Mean language score§	2.4	2.4	2.0	1.9*
Mean days of alcohol use††	0.9	1.0	0.9	0.9
Mean days per week home alone	2.4	2.4	2.7	2.8
Mother’s mean educational level‡‡	2.8	2.8	3.2	3.4
% who ever had a serious boyfriend/girlfriend	35.2	36.3	39.1	37.9
% who ever had sex	9.7	10.9	11.0	10.0
Mean no. of sexual partners who were ever pregnant	0.2	0.2	0.3	0.2
% who were ever diagnosed with STD	0.4	0.2	0.5	0.4
	0.3	0.0	0.3	0.3
Sample range	605–893		2,933–3,396	

*Difference between treatment and control group is significant at $p < .01$. †Ethnicity percentages do not add to 100% because of missing data, and because multiracial youths could check more than one category. ‡Academic grades were scored as 1=mostly As and Bs and 4=mostly Ds and Fs. §Language score is the degree to which English was spoken at home and was scored as 1=mostly English and 4=mostly other language. ††Indicates number of days within the last month that respondent drank alcohol. ‡‡Mother’s educational level was scored as 1=eighth grade or less and 5=college graduate.

apparent at 17 months. Although the impact of the youth-led intervention was almost as large as that of the adult-led group, it did not reach statistical significance.

There were no statistically significant differences between the treatment and control groups in the number of reasons or conditions under which youths said they would engage in sex, at either three or 17 months. For youths who had not had sex at pretest, the pattern of statistical significance was the same, and the effect size was similar (not shown).

• *Beliefs about sex and the media.* At the three-month posttest, youths who participated in either the adult-led or the youth-led PSI program were significantly more likely than teenagers in the corresponding control groups to recognize the sex-related content of media messages (effect

*To estimate the size of this difference based on a standard metric, we computed effect sizes using a pooled standard deviation. Generally, effect sizes of less than 0.20 are considered small and those greater than 0.50 are considered substantial (see: J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed., Academic Press, New York, 1988).

Table 2. Mean change in selected mediating variables among students in treatment and control groups, by type of intervention and setting, according to length of follow-up

Variable	Youth-led				Adult-led											
	Classroom				Classroom				School				Community			
	Treat-ment	Con-trol	p	N	Treat-ment	Con-trol	p	N	Treat-ment	Con-trol	p	N	Treat-ment	Con-trol	p	N
BELIEFS ABOUT SEXUAL ACTIVITY																
Postponing sex																
3 months	-.038	-.119	.003*	1,668	-.050	-.106	.011	2,363	na	na	na	na	na	na	na	na
17 months	-.237	-.274	.308	1,363	-.251	-.282	.273	2,038	-.323	-.315	.710	3,297	-.103	-.105	.983	338
Inevitability of teenage sex																
3 months	.063	-.056	.006*	1,805	.096	-.040	.000*	2,581	na	na	na	na	na	na	na	na
17 months	-.149	-.132	.754	1,470	-.044	-.135	.027	2,208	-.083	-.102	.572	3,651	.042	-.058	.369	362
Sexual pressure																
3 months	.004	-.010	.606	1,689	.027	-.022	.027	2,428	na	na	na	na	na	na	na	na
17 months	-.077	-.041	.312	1,388	-.015	-.041	.322	2,103	.001	-.019	.348	3,489	.013	.140	.053	341
Peers' sexual activity																
3 months	-.132	-.133	.980	1,731	-.103	-.128	.517	2,481	na	na	na	na	na	na	na	na
17 months	-.773	-.651	.078	1,419	-.716	-.660	.294	2,138	-.704	-.711	.865	3,612	-.527	-.605	.572	357
Possible to say no to sex																
3 months	.053	-.048	.033	1,803	.030	-.017	.227	2,595	na	na	na	na	na	na	na	na
17 months	-.028	-.075	.376	1,475	.020	-.054	.094	2,224	.054	.011	.219	3,696	.037	-.098	.234	360
REASONS TO HAVE SEX OR ABSTAIN																
Not to have sex																
3 months	-.008	-.059	.011	1,592	.010	-.047	.000*	2,308	na	na	na	na	na	na	na	na
17 months	-.163	-.141	.418	1,307	-.094	-.132	.083	1,985	-.134	-.123	.487	3,454	-.027	.023	.344	339
To have sex																
3 months	-.023	-.053	.112	1,629	-.030	-.044	.401	2,350	na	na	na	na	na	na	na	na
17 months	-.110	-.114	.873	1,332	-.107	-.108	.992	2,032	-.118	-.103	.354	3,552	.041	.049	.898	343
BELIEFS ABOUT SEX AND THE MEDIA																
Media encourage sex																
3 months	.194	.022	.001*	1,791	.263	.048	.000*	2,571	na	na	na	na	na	na	na	na
17 months	.173	.094	.199	1,459	.194	.079	.016	2,196	.149	.090	.108	3,673	.011	.162	.152	362
Media do not affect behavior																
3 months	.110	.009	.047	1,775	.122	.020	.016	2,548	na	na	na	na	na	na	na	na
17 months	.054	-.033	.158	1,453	.013	-.011	.627	2,185	-.021	-.020	.976	3,659	.131	.132	.991	365
Sex is used to sell products																
3 months	.186	.005	.000*	1,791	.113	.020	.004*	2,559	na	na	na	na	na	na	na	na
17 months	.175	.103	.145	1,470	.083	.105	.568	2,190	.130	.061	.020	3,709	.083	.253	.083	366

*Difference between treatment and control group is significant. Notes: In this and subsequent tables: Larger mean change scores represent more desirable outcomes. The treatment group in the school setting received PSI along with schoolwide ENABL activities. na=not applicable, because three-month follow-up occurred only in classroom setting.

sizes of 0.11–0.21). No statistically significant differences remained at the 17-month follow-up. Among youths who had not had sex at pretest, the patterns of statistical significance were the same, but the effect sizes were slightly larger (0.14–0.24). There were no differences between treatment and control groups at either three or 17 months in youths' belief that the media have no influence on their behavior.

• *Communication with parents.* Regardless of sexual experience at baseline, there were no statistically significant differences at either the three- or 17-month follow-up between treatment and control groups in the level of communication with parents during the preceding year (not shown).

• *Self-efficacy.* As Table 3 indicates, youths receiving the adult-led but not the youth-led PSI were significantly more likely at the three-month posttest than youths in the control group to have confidence that they

could say no to sex. The intervention appears to have counteracted a maturation effect, as youths who participated became more likely to believe they could refuse sex, while those in the control groups became less likely to believe so. However, the difference between the adult-led intervention and the corresponding control group was small (0.09, with an effect size of 0.13), and it did not remain significant at 17 months. There were no significant differences between treatment and control groups, at either three or 17 months, in participants' belief that they could demonstrate affection without having sex. The pattern of statistical significance was the same among youths who had never had sex at pretest (not shown).

• *Behavioral intentions.* There were no statistically significant differences between treatment and control groups in youths' having set sexual limits, in their intention

to avoid sex even at the risk of losing a relationship (not shown), in sexually inexperienced youths' deciding to wait until they are older to have sex, or in sexually experienced youths' deciding to refrain from sex in the near future.

At the three-month posttest, teenagers in the youth-led but not the adult-led intervention were significantly more likely than control youths to report intending to refuse sex even when stirred by sexual feelings (effect size of 0.12), and those in the adult-led but not the youth-led intervention were significantly more likely than control youths to indicate that they intended to refuse pressure to have sex (effect size of 0.13).

Sexual Behavior

There were no statistically significant differences in attempts to persuade others to engage in intercourse between treatment and control groups in any setting at either

Table 3. Mean change in self-efficacy and in behavioral intentions among students in treatment and control groups, by type of intervention and setting, according to length of follow-up

Variable	Youth-led				Adult-led											
	Classroom				Classroom				School				Community			
	Treat-ment	Con-trol	p	N	Treat-ment	Con-trol	p	N	Treat-ment	Con-trol	p	N	Treat-ment	Con-trol	p	N
SELF-EFFICACY																
Saying no to sex																
3 months	.011	-.034	.218	1,537	.046	-.048	.003*	2,243	na	na	na	na	na	na	na	na
17 months	-.115	-.035	.100	1,265	.008	-.054	.116	1,939	-.046	-.061	.609	3,494	.048	.080	.738	334
Affection without sex																
3 months	-.026	-.065	.357	1,825	-.029	-.039	.756	2,617	na	na	na	na	na	na	na	na
17 months	-.023	-.037	.774	1,504	-.005	-.056	.193	2,254	-.017	-.026	.764	3,708	.047	.082	.738	361
BEHAVIORAL INTENTIONS																
Have set sexual limits																
3 months	.057	.042	.607	1,584	.027	.062	.132	2,289	na	na	na	na	na	na	na	na
17 months	.089	.105	.645	1,264	.108	.105	.900	1,967	.107	.095	.579	3,364	.052	.075	.738	332
Sexual feelings would lead to sex																
3 months	-.038	-.160	.007*	1,910	-.070	-.129	.120	2,731	na	na	na	na	na	na	na	na
17 months	-.317	-.270	.440	1,574	-.215	-.245	.534	2,339	-.299	-.277	.565	3,823	-.117	-.147	.801	381
Sexual pressure would lead to sex																
3 months	-.089	-.184	.026	1,918	-.051	-.165	.001*	2,739	na	na	na	na	na	na	na	na
17 months	-.310	-.244	.242	1,577	-.246	-.229	.697	2,348	-.299	-.255	.201	3,825	-.036	-.151	.293	382
Sexual pressure from boyfriend/girlfriend would not lead to sex																
3 months	-.053	-.021	.580	1,599	.014	-.036	.301	2,296	na	na	na	na	na	na	na	na
17 months	-.122	-.008	.127	1,328	-.063	.004	.249	1,979	-.098	-.082	.722	3,387	.087	-.073	.221	337
Have decided to postpone sex†																
% at 17 months	62.2	63.0	.768	1,197	66.4	63.1	.150	1,782	61.3	63.4	.253	2,829	71.3	67.5	.457	330
Have decided to postpone further sex‡																
% at 17 months	23.1	20.6	.551	396	20.9	22.6	.641	590	21.3	18.6	.288	1,016	34.5	33.3	.930	53

*Difference between treatment and control group is significant. †Among all youths not sexually experienced at 17-month follow-up. ‡Among youths sexually experienced at 17-month follow-up.

three or 17 months (not shown). As Table 4 (page 106) indicates, among youths who reported never having had intercourse at baseline, there were no statistically significant differences between intervention and control youths in the percentage who had initiated intercourse at either the three-month (5–6%) or the 17-month follow-up (15–18% in school settings and 8% in community settings). Furthermore, we found no significant differences in the impact of PSI on the postponement of sexual intercourse for different subgroups of students according to gender, grade, race or ethnicity, history of serious romantic involvement, prior receipt of sex education or contract agency responsible for implementing the program (not shown).

Among sexually experienced youths, there were no significant differences at either follow-up point in frequency of intercourse or number of sexual partners between any of the treatment and control groups, regardless of implementation setting, age of group leader (youth or adult) or participant's gender, prior sexual history, grade level, race or ethnicity or sexual experience prior to baseline.

We also examined the possibility that because the program discussed pregnan-

cy and STDs as consequences of sexual activity and because it taught assertiveness skills to avoid sex, it might help adolescents insist upon the use of contraceptives if they did have sex. However, no significant differences emerged between intervention and control youths' use of oral contraceptives or condoms.

Pregnancy and STDs

Data on pregnancy rates among all youths who reported no prior history of pregnancy at pretest are presented in Table 5 (page 107) only for the 17-month follow-up, since any pregnancies reported at the three-month posttest were likely to have been conceived prior to baseline. There were no statistically significant differences in pregnancy rates between teenagers receiving the adult-led intervention and those in the corresponding control groups. This was true for all three settings, when analyzed separately or in combination. Furthermore, there were no significant differences when data only from those youths in adult-led groups who were sexually experienced at baseline were analyzed.

However, there was a statistically significant difference among those who had received the youth-led intervention; they

were more likely to report a pregnancy than were their control group counterparts (4% vs. 2%). The same pattern emerged when only those youths sexually experienced at baseline were analyzed. We then conducted a multilevel statistical analysis, adjusting for clustering of youths within classrooms as well as within schools; the results remained statistically significant.

Given our other findings, this significant result was highly unexpected; we completed several additional analyses to more fully understand its cause. Analysis by gender revealed that the major difference between the treatment and control groups was among males. Young men receiving youth-led PSI reported a remarkably high rate of pregnancy involvement compared with their counterparts in the control group (6% vs. 2%). Further analyses revealed that a disproportionate number of the males reporting involvement in a pregnancy came from the seventh grade class of one particular school: Six males in the intervention group and only one in the control group reported having caused a pregnancy. When all seventh graders from that one school were removed from the statistical analysis, the overall relationship between youth-led PSI and pregnancy was

Table 4. Measures of sexual and contraceptive behavior among students in treatment and control groups, by type of intervention and setting, according to sexual experience and length of follow-up

Variable	Youth-led				Adult-led											
	Classroom				Classroom				School				Community			
	Treat-ment	Con-trol	p	N	Treat-ment	Con-trol	p	N	Treat-ment	Con-trol	p	N	Treat-ment	Con-trol	p	N
SEXUALLY INEXPERIENCED AT BASELINE																
% who initiated intercourse																
At 3 months	5.7	5.2	.673	1,678	5.6	5.4	.800	2,381	na	na	na	na	na	na	na	na
At 17 months	16.7	15.8	.660	1,431	15.1	17.1	.210	2,134	18.4	16.5	.126	3,446	7.7	8.4	.798	362
SEXUALLY EXPERIENCED AT EITHER FOLLOW-UP																
No. of acts of intercourse in last 3 months†																
Mean change at 3 months	0.9	0.3	.023	338	0.8	0.4	.109	464	na	na	na	na	na	na	na	na
Mean change at 17 months	1.8	1.6	.619	338	2.0	1.6	.162	586	1.7	1.9	.531	1,012	2.0	1.9	.960	52
No. of acts of intercourse in last 12 months†,‡																
Mean change at 17 months	3.0	2.7	.454	389	2.9	2.7	.453	586	2.8	3.0	.340	1,010	2.9	2.3	.552	52
No. of sexual partners ever†																
Mean change at 3 months	1.1	0.7	.069	342	1.0	0.8	.120	470	na	na	na	na	na	na	na	na
Mean change at 17 months	2.3	2.0	.282	393	1.9	1.8	.643	584	1.9	2.0	.420	1,012	1.0	1.4	.596	53
Used a condom																
% at 3 months	59.7	56.8	.593	339	65.9	56.3	.034	471	na	na	na	na	na	na	na	na
% at 17 months	56.7	61.0	.394	394	61.4	60.4	.802	584	66.2	67.0	.792	1,012	62.1	66.7	.728	53
Used oral contraceptives																
% at 3 months	24.3	20.4	.385	343	31.6	22.4	.024	475	na	na	na	na	na	na	na	na
% at 17 months	21.4	26.9	.212	387	19.4	24.3	.160	585	23.3	21.4	.469	1,015	20.7	29.2	.475	53

†Youths not sexually experienced at baseline were assigned a value of 0. ‡We do not include three-month data for this variable, because this time period included the nine months prior to baseline.

no longer statistically significant (p=.14).

There are a variety of possible explanations for what happened among those six seventh grade males. Given that classrooms of students were randomly assigned, it is possible that an especially high-risk group of males may have been assigned to one classroom, which was then assigned to the treatment condition. Alternatively, a small cluster of males may have decided to report incorrectly that they had caused a pregnancy, there may have been some gang activity requiring sexual activity and the claim of paternity, or several males may have each thought they were responsible for a single pregnancy.

In any case, it is unlikely that PSI caused an actual increase in pregnancy rates, for several reasons: First, much of the difference in pregnancy rates between students in the youth-led PSI and those in the corresponding control group occurred only among males in the seventh grade class in one school and did not occur in other classes, schools or agencies, or among females. Additionally, the males receiving youth-led PSI reported extremely high rates of pregnancy in relation to statewide statistics. Finally, there were no significant differences between treatment and control groups in sexual behavior and contraceptive use that would explain differences in pregnancy rates. Thus, the weight of the evidence indicates that neither youth-led nor adult-led PSI had a signifi-

cant effect upon actual pregnancy rates.

Of those students who at pretest had reported never having had an STD, there were no significant differences at either follow-up between the PSI groups—either youth-led or adult-led—and their respective control groups in the percentage of youths who reported an STD (Table 5). Furthermore, there were no significant differences between the youth-led group and their control counterparts when the analysis was restricted to only sexually experienced youths.

There were no significant differences in reported STD rates between intervention and control participants in the schoolwide randomization or among teenagers recruited from community settings. (Rates among intervention groups were lower, although not significantly so.) However, participants in adult-led intervention groups in the classroom randomization design had significantly higher STD rates than did their control group counterparts.

Reported STD rates can increase either because greater proportions of youths actually contracted an STD or because larger proportions of youths with an STD decided to be tested and therefore learned that they had an STD. Thus, it is not clear whether an increase in reported STD rates represents a desirable or an undesirable event. Moreover, two of the five STDs listed in the survey question (herpes and crab lice) can be transmitted without sexual intercourse.

It is likely that this significant finding occurred by chance. The STD rates for the intervention and control groups in all three adult-led settings combined were remarkably similar (5–6%), while in the two adult-led settings other than the classroom randomization design, youths in the PSI groups had lower STD rates than did those in the control groups. Moreover, youths in the classroom scheme who participated in the adult-led intervention had not had significantly higher rates of sexual intercourse than their control group counterparts during the previous year (2.8 vs. 2.9 sexual acts), nor did they have a significantly higher number of sexual partners (1.9 vs. 1.8), nor were they significantly less likely to use condoms the last time they had sex (61% vs. 60%). Thus, there is no causal explanation for this finding.

However, it is possible that these students were more likely to obtain STD testing; some PSI leaders gave students referral cards specifying where such testing could be obtained. The weight of our findings strongly suggests that the PSI intervention did not significantly affect actual rates of STD infection.

Discussion

These results provide a remarkably consistent picture of the impact of PSI and the ENABL initiative. In the short term, the intervention had no impact on seven beliefs and attitudes, on four measures of inten-

Table 5. Percentage of youths reporting a pregnancy (or causing a pregnancy) or a diagnosed STD, by type of intervention and setting, according to length of follow-up

Variable	Youth-led				Adult-led											
	Classroom				Classroom				School				Community			
	Treat-ment	Con-trol	p	N	Treat-ment	Con-trol	p	N	Treat-ment	Con-trol	p	N	Treat-ment	Con-trol	p	N
Ever pregnant or caused a pregnancy†,‡																
At 17 months	3.7	1.5	.006*	1,548	2.5	1.9	.265	2,250	3.0	2.2	.139	3,758	1.1	1.1	.970	371
Ever diagnosed with an STD§																
All respondents																
At 3 months	0.7	0.3	.195	1,895	1.0	0.3	.032	2,711	na	na	na	na	na	na	na	na
At 17 months	1.5	0.7	.133	1,545	1.6	0.6	.015	2,313	1.4	1.6	.684	3,761	0.5	1.7	.289	372
Sexually experienced respondents																
At 3 months	4.0	1.1	.080	334	6.4	1.6	.007*	460	na	na	na	na	na	na	na	na
At 17 months	5.9	2.7	.122	389	6.8	2.1	.005*	584	5.1	6.7	.306	979	3.8	13.6	.221	48

*Difference between treatment and control is significant. †Among youths never pregnant at baseline. ‡We do not include three-month data for this variable, because pregnancies during this time period could have been conceived prior to baseline. §Based only on those who at baseline never had a diagnosed STD.

tions to have sex, or on five measures of sexual behavior.

The intervention had a small, positive impact among some groups on several attitudes related to sexual decision-making, on perceptions about the media's presentation of sexual images, and on feelings of self-efficacy and intentions to refuse sex. These attitudinal shifts did not translate into positive behavioral changes. Moreover, at 17 months, the intervention had no significant and positive effect upon any mediating variable, upon sexual or contraceptive outcomes or upon pregnancy or STD rates.

Our findings raise an important question: Why did this evaluation reveal no behavioral impact at three months and no impact of any kind at 17 months, when the evaluation of PSI in Atlanta suggested behavioral change? Is it possible that we failed to detect significant positive outcomes? The answer to this question has important implications for recommendations about how to develop effective programs.

Strengths and Limitations

This evaluation had several strengths: It employed a strong design with random assignment, short- and long-term follow-up and large sample sizes with sufficient statistical power to detect programmatically meaningful effects. It also allowed for the evaluation of youth-led PSI and adult-led PSI in schools and adult-led PSI in community settings. Moreover, this evaluation accounted for most of the mediating variables that might be affected by educational interventions and, in turn, might affect the initiation of intercourse. Most of the scales employed had acceptable-to-high reliability, and most of the behavioral measures had high internal consistency, both within each survey and between surveys. Finally, we checked ex-

tensively for inconsistencies and removed individuals with discrepant data.

However, several limitations are also noteworthy. This study did not have a strict no-treatment control group. While youths in the control groups received whatever program or instruction was otherwise being offered, it typically did not cover human sexuality. In addition, a large majority of the youths in both the treatment and control groups in this study, like those in Atlanta, had previously received some other instruction about aspects of human sexuality at some time during their middle school years. Thus, we could not assess whether PSI was more effective than nothing; all we could evaluate was whether PSI had a significant impact when it was taught in addition to other limited instruction on human sexuality.

There are several measurement limitations that are also noteworthy. Although youths who overreported or underreported sexual activity are likely to have been randomly distributed between the intervention and control groups, some youths who reacted negatively to the program or who were rebelling against its messages may have disproportionately overreported their sexual behavior. On the other hand, it is possible that youths who participated in PSI began to see teenage sexual activity in a less favorable light, and consequently underreported their own sexual activity. In either case, however, it seems likely that youths would overreport or underreport at the three-month posttest rather than at the 17-month follow-up, when any program effects are likely to have diminished. Moreover, our data on rates of sexual behavior are consistent with those from other studies.

The measurement of pregnancy is somewhat problematic. Several weeks or

even several months may elapse between conception and the time a young woman receives results from a pregnancy test. Until that time, she may not know of a conception or may incorrectly believe she is pregnant when she is not. Males, on the other hand, may not know that they have caused a pregnancy unless their sexual partner tells them.

Overall, we feel it is unlikely that the interventions produced programmatically important effects that were not detected. In the context of a strong design and methodology, we examined many subgroups of youths and searched at length for significant, positive and consistent behavioral effects. We found insufficient change in the mediating variables to suggest that there could be significant change in behavioral outcomes, and the results were remarkably consistent in demonstrating that PSI did not produce desirable effects upon behavior. Finally, behavioral results frequently were not in the desired direction, were not programmatically significant and were not close to statistical significance.

Replication Issues

When programs are replicated and implemented broadly, they are not always replicated with high fidelity. Accordingly, the ways in which the California implementation differed from that of the Atlanta evaluation should be examined.

The scale of the implementation in California was dramatically larger than that in Atlanta, and contractors had to stretch their resources and capacities in order to deliver PSI to large numbers of youths in relatively short periods of time. This raises the possibility that elements of the program may not have been implemented with the same fidelity as in Atlanta. There are several ways in which the California implementation of

PSI differed from that in Atlanta.

• *Age of students.* In Atlanta, only eighth grade students participated in the PSI evaluation, whereas in California, the program involved both seventh and eighth graders. However, when eighth graders were analyzed separately in California, results were similar to the findings from the combined analysis.

• *Additional five-session unit.* In Atlanta, PSI was implemented in addition to a five-session reproductive health unit that included basic human sexuality, decision-making and contraception, and the evaluation actually measured the impact of both PSI and this five-session reproductive health unit. In the California replication, the state contract required that study participants receive reproductive health education prior to receiving PSI, and 85% of our sample specifically remembered receiving such instruction. However, this instruction did not necessarily occur immediately prior to PSI, as it did in the Atlanta implementation. Even so, findings based upon only those youths who remembered that they had previously received this instruction did not differ significantly from those based upon all youths. In addition, it does not seem likely that a series of classes that focused on postponing sex would be ineffective in delaying the onset of intercourse, yet would be successful in doing so if additional information were added on reproductive health and contraception.

• *Group leaders.* The PSI curriculum was developed for implementation by teenagers, and this is how it was implemented in Atlanta. In the California replication, the program was largely implemented by adults. Although this is an important difference, the results of this study were not more positive for those teenagers who participated in the youth-led intervention than for those who participated in the adult-led intervention.

• *Video.* The PSI curriculum came with a video showing still photographs of youths accompanied by voiceover narrations. About half of the ENABL project contractors used this video, but the rest found that the youths they served reacted so unfavorably to it that they could not use it. Thus, many of the youths in this evaluation received PSI without the video.

• *Implementation.* The educators implementing PSI—both youths and adults—were specially trained to deliver the program, were contractually obligated to follow the curriculum and knew they were being evaluated. Moreover, many

contractors assigned their best and most experienced educators to facilitate the groups being evaluated. Thus, it is likely that in most respects, the basic structure and activities of the PSI curriculum were followed closely. Only a few modifications affecting only a small proportion of study participants were approved, and these were designed to make the curriculum more culturally appropriate.

Our personal observations and reports from agencies confirm that most contractors did follow the curriculum with considerable faithfulness. Moreover, the curriculum is well scripted, and according to most sexuality educators, relatively easy to follow. Group leaders received two days of training and practice in how to implement the program, and most belonged to organizations that commonly deal with sexual issues (e.g., family planning agencies); most had taught sex education in the classroom.

However, our personal observations of classroom instruction indicated that not all of the adult leaders always gave sufficient emphasis to important program messages; they sometimes spent more time than necessary answering questions not directly related to the intervention's goals. Moreover, a few leaders expressed dissatisfaction with the intervention's primary focus on postponing sexual involvement and the exclusion of information about contraception and disease prevention. Their conflicting feelings about the program may have diluted the strength of the messages they presented to students. Thus, some of the leaders may not have implemented PSI with optimal clarity and skill.

It was also our observation that some of the teenagers who led the intervention groups were not sufficiently trained or experienced. Some of the youths were not entirely comfortable talking about sex or communicating the program's singular message about postponing sexual involvement.

Conclusions

Despite these concerns, we do not believe that simply improving the fidelity of the implementation will cause PSI to dramatically change sexual behaviors. Rather, we believe that although the development of PSI was a seminal event in our field and the curriculum has broad appeal, the intervention (at five sessions in length) is too modest to have a significant impact on behavior. Indeed, the only three curricula implemented in the classroom that have led to changes in adolescent sexual behavior

lasted an average of 15 sessions.⁷

Furthermore, the PSI program lacks one essential element of a successful behavior change curriculum: the opportunity to learn and practice new skills within an environment that provides sufficient support and feedback. Given its modest length, PSI cannot provide much practice in skill-building; during some implementations, a few participants did not have even a single opportunity to practice a refusal.

There is currently no middle school curriculum for which strong evidence indicates it is effective in delaying sexual involvement among young adolescents. Thus, there remains a real need to develop and demonstrate the effectiveness of such a program.

The findings from this replication study also make clear that before any group broadly implements a specific curriculum, it should thoroughly and critically examine the evidence for the effectiveness of that curriculum. Such a review should consider whether the curriculum was implemented, evaluated and found to be effective in large and rigorous studies in multiple sites. Characteristics of individual sites—uniqueness of the target population, unusually charismatic leaders or vagaries of the evaluation design—may limit the generalizability of the findings from one site to other sites. Thus, positive findings should be demonstrated in multiple sites and preferably multiple studies before a program is broadly replicated.

References

1. M. Howard and M. E. Mitchell, *Postponing Sexual Involvement: An Educational Series for Young Teens*, Emory/Grady Teen Services Program, Grady Memorial Hospital, Atlanta, Ga., 1990.
2. M. Howard and J. McCabe, "Helping Teenagers Postpone Sexual Involvement," *Family Planning Perspectives*, 22:21–26, 1990.
3. D. Kirby, *A Review of Educational Programs Designed to Reduce Sexual Risk-Taking Behaviors Among School-Aged Youth in the United States*, National Technical Information Service #PB96108519, Springfield, Va., 1995.
4. H. H. Cagampang et al., "Education Now and Babies Later (ENABL): Life History of a Campaign to Postpone Sexual Involvement," *Family Planning Perspectives*, 29:109–114, 1997.
5. M. Howard and J. McCabe, 1990, op. cit. (see reference 2).
6. D. Kirby et al., *Evaluation of Education Now and Babies Later (ENABL): Final Report*, Family Welfare Research Group, Berkeley, Calif., 1995.
7. S. Schinke, B. Blythe and L. Gilchrist, "Cognitive-Behavioral Prevention of Adolescent Pregnancy," *Journal of Counseling Psychology*, 28:451–454, 1981; D. Kirby et al., "Reducing the Risk: A New Curriculum to Prevent Sexual Risk-Taking," *Family Planning Perspectives*, 23:253–263, 1991; and D. S. Main et al., "Preventing HIV Infection Among Adolescents: Evaluation of a School-Based Education Program," *Preventive Medicine*, 23:409–417, 1994.