An Evaluation of the Heritage Keepers® Abstinence Education Program

Stan E. Weed, Ph.D., Irene H. Ericksen, M.S. and Paul James Birch, M.S.

Institute for Research and Evaluation, Salt Lake City, UT

Abstract

An abstinence education program for middle school students in South Carolina was evaluated using a quasi-experimental design with matched comparison groups to determine its impact on the rate of sexual initiation after 12 months for the sexually inactive in the sample. Pretest, posttest, and 1-year follow-up measures were taken of psychosocial factors (known predictors or mediators of adolescent sexual initiation) and of the sexual behavior of the sample. These three measures were linked for individual students to allow for a repeated measures analysis. A predictive model of sexual initiation was confirmed using the mediator variables. The 12-month follow-up measure indicated a significant reduction in sexual initiation for the program participants (OR=.54, $p<.001$), after controlling statistically for some pretest differences between program and comparison groups. Confirmatory analyses done to further test the impact of pretest differences and attrition at follow-up consistently produced odds ratios in the range of .44 to .50 for the sexual initiation of the program versus comparison group. The conclusion that there was a significant and sustained program effect on initiation rates seems warranted.

Introduction

Primary prevention strategies for adolescent risk behavior have become increasingly prevalent in recent years as a method to ameliorate the results of premature sexual activity. The negative consequences of teen pregnancy are well documented (Miller and Moore 1990; Maynard 1997; Jaffee 2002). Sexually transmitted diseases are a growing threat to adolescent health because of their high prevalence in this age group, their destructive results, and the difficulty of preventing some STDs by means of barrier methods (Centers for Disease Control 2001; National Institute of Allergy and Infectious Diseases 2001; Sulack 2003; Genuis and Genuis 2004; Weinstock, Berman, and Cates 2004).

Primary prevention—risk avoidance

In its 2000 issuance of *Healthy People 2010*, the U.S. Department of Health and Human Services included the following in its list of protective behaviors that are its first line of defense for the problems attendant to teen sexual activity:

> "The protective behaviors of interest are completely abstaining from sexual intercourse during adolescence (primary abstinence), [and] reverting to abstinence for long periods of time after having had intercourse in the past (secondary abstinence)..."

The key question is whether or not adolescents can be influenced to abstain from sexual activity. Over the past decade the proportion of U.S. teens who have not engaged in sexual intercourse has been growing. Presently more than half (53.3 percent) of high school students are in this category (Grunbaum and others 2004). In addition, some 82 percent of U.S. teens said that teens should not be sexually active, according to the National Campaign to Prevent Teen Pregnancy's 2003 national survey. Among those who had already initiated sexual activity, 67 percent, in the same survey, said they wish they had waited. These statistics suggest that it may be possible to influence teens towards sexual abstinence.

Increased interest in primary prevention has led to a proliferation during the past decade of abstinence education programs aimed at American teens. However, efforts to evaluate such programs up to this point in time have been limited—in number and in rigor (Thomas 2000; Kirby 2001; Manlove and others 2002; Scher and others 2005). For the few evaluations that are available, there is often disagreement about the quality of the research and the meaning of the results (Rector 2002; Kirby 2002b). There is agreement, however, that more and better program evaluations need to be done. The present study attempts to address this need.

A risk avoidance model for adolescent sexual behavior

The lack of research about effective abstinence interventions has been compounded by a lack of understanding of the causal mechanisms that influence teen abstinence. The developers of Heritage Keepers® Abstinence Education sought to identify causal mechanisms that would influence teen sexual abstinence, and then develop a model around which to design their abstinence program. Psychosocial constructs are an important focus in the study of behavior change, in part because they are strong predictors of behavior and are more amenable to manipulation than demographic or environmental factors (Armitage and Conner 2000). The predictive power of these variables has held up well in empirical studies of health behavior (Armitage and Conner 2000; Strecher and others 1986; Godin and Kok 1996; Floyd, Prentice-Dunn, and Rogers 2000). In addition to their established role in health behavior models, psychosocial factors have been found to be important in understanding adolescent sexual behavior (Miller and Moore 1990; Kirby 2001; Plotnik 1992; Resnick and others 1997; Kirby 2002a). Preliminary work by Weed and Olsen (1988) and Weed and others (1992) explored a broad set of cognitive/affective constructs as predictors of sexual initiation versus abstinence for several samples of U.S. teens.

From this work, in conjunction with the literature on antecedents of teen sexual behavior (Miller and Moore 1990; Plotnik 1992; Resnick and others 1997; Kirby 2002a) program developers formulated a predictive model with features similar to the theory of planned behavior, social cognitive theory, and protection motivation theory (Armitage and Conner 2000; Floyd, Prentice-Dunn, and Rogers 2000; Bandura 2004; Ajzen 1991; Conner and Armitage 1998). This model posits *behavioral intentions* as the primary antecedent or first-order predictor of teen sexual abstinence (as measured by rates of sexual

initiation). Variables predicting behavioral intentions (second-order predictors) are in these categories: self-efficacy, which they label *abstinence efficacy*; outcome expectancies called *future impact of sex*; attitudes or *justifications for sex*; and moral norms, called *abstinence values*. This model was incorporated into the Heritage curriculum design, to be tested as part of the program evaluation.

Purpose of the study

The ultimate goal of the Heritage program is primary prevention, i.e., to influence adolescents to postpone sexual initiation until marriage. Therefore the focus of this evaluation was the program's impact on sexually inactive youth. The primary purpose of the study was to evaluate the impact of an abstinence intervention on both the mediating variables and the behavioral outcome of sexual initiation for the sexually inactive students in the sample. A secondary purpose was to examine further the relationship between the mediating variables and the behavioral outcome of teen sexual initiation. A third purpose was to assess the program's system for continuous improvement, i.e., for monitoring and maintaining a high level of implementation and fidelity to program design and for optimizing teacher effectiveness.

Four central hypotheses were tested to accomplish these purposes.

- There will be a direct and significant relationship between the second-order, mediating variables *abstinence values, justifications for sex, abstinence efficacy, and future impact* and the first-order mediator *behavioral intention.*

- There will be a direct and significant relationship between the mediator *behavioral intention* and the behavioral outcome of *sexual initiation.*

- Program group students will demonstrate significant pre–post improvement on the mediating variables, and this improvement will be significantly greater than the pre–post change for the comparison group. The program group will maintain a position that is significantly better on the mediating variables than the comparison group at the 12-month follow-up.

- After 12 months, virgin program students will be significantly less likely to have initiated sex than the virgin comparison group students.

Program description

Heritage Keepers® Abstinence Education is a 450-minute, interactive curriculum that is designed for middle and/or high schools. It is presented in 45-minute class periods over 10 consecutive school days or in 90-minute sessions for five consecutive days. This level of annual program dosage is intended for presentation to students over three successive years. There are two levels of curriculum, one that establishes a foundation and a second that builds upon and expands the concepts that were presented in the first level. The first level is to be presented in each of two consecutive years and the second level curriculum is intended for the third year. The curriculum content is based on the Title V, Section 510 (b)(2)(A)–(H) guidelines, with particular attention to the mediating constructs identified as causal mechanisms for influencing teen sexual behavior. These concepts are delivered by teachers who are selected and trained by Heritage to relate well with students and engage them in active learning processes that foster commitment to abstinence.

Method

Sample

Heritage Keepers® promotes its abstinence program on a regional basis in South Carolina through outreach directly to middle and high schools. Program schools contain the range of demographic profiles and risk levels found among students in the state. The program schools used in this evaluation were the subset of the Heritage schools where fidelity to plan was considered high. Comparison schools were selected based on demographic and geographic similarity to these schools, as well as the fact that they had implemented no abstinence education program. This evaluation was approved by the school board for each district in the sample, and parental approval was gained through passive consent.

The sample for this study consisted of students in grades 7–9 from 34 program schools and 7 comparison schools in South Carolina. The Heritage curriculum was presented in required classes at each school. All students in these classes participated in the program unless their parents exempted them from the program. The rate of parental refusal was low, at approximately 3 percent.

There were 2,529 virgin students in the program schools and 417 in the comparison schools. The number of virgin students whose 12-month follow-up surveys could be linked with their pretest survey (using a confidential ID code) was 1,281 for the program group and 254 for the comparison groups, for a retention rate after one year of 50.7 percent and 60.9 percent respectively. This level of attrition seems consistent with the unusually high student mobility and dropout rates for South Carolina, which one ranking showed having the lowest high school promotion rate in the country at 51 percent (Hall 2005). In our study, this was compounded by a scheduling conflict with state-mandated testing that caused several hundred follow-up cases to be unobtainable. Although the percent retained was lower for the program group than the comparison group ($p < .001$), the impact of this attrition difference on group comparability was negligible. For both groups, those lost to attrition were somewhat higher in risk propensity, but the attrition differences between groups were insignificant, so their relative distribution on grade, race, and mediator scores remained about the same after attrition as before. The main concern is whether the pretest program and comparison groups are similar on measures of propensity for risk behavior.

Table 1 presents a description of the linked sample, with significant differences between program and comparison groups indicated. The demographic composition and pretest scores on mediating variables made the comparison group appear to be a somewhat higher-risk sample. These pretest differences were considered to be in a range within which adjustments could be made statistically to compensate for any impact on outcomes. The analyses that adjust for these differences are described in the results and discussion section.

Research design

The evaluation study used a quasi-experimental design with matched comparison schools, repeated measures, and one year follow-up. Pre- and posttest surveys were conducted for two different cohort years, with a 12-month, follow-up survey for each cohort. In a few schools in our program sample, the Heritage curriculum had operated the year prior to our initial data collection, meaning that students in the older grades probably received at least one dose of the program prior to taking the pretest. (The reliability of self-report on prior program exposure was not adequate for classifying students.) Regardless of prior exposure, all pretest students were followed for one year, to assess program effects on behavior over time. The 1-year follow-up measure was taken prior to the commencement of the curriculum installment for that next year. (Where prior

program exposure occurred, the effect would likely be to minimize the measures of positive program impact.)

Posttest surveys were administered by the abstinence teacher in the classroom. One-year follow-up surveys were typically conducted by a different teacher than the one who taught the students the previous year. Questionnaires were anonymous; there were no identifying marks and the students deposited their surveys directly into a box. Identification codes were created later for linking purposes but never connected to the student's name. Prior to administering the questionnaire, teachers reviewed these anonymity procedures with students, encouraged them to be candid, and reminded them of the importance of accurate responses.

Instrumentation

A 58-question survey was administered at the pre, post, and 12-month follow-up. It included demographic questions, behavioral questions, and multiple items that formed five scaled measures of the mediating constructs. The key demographic questions included gender, ethnicity, grade, and repeated grade. The behavioral questions included a basic question of whether or not students had ever had sexual relations, with follow-up questions, such as age of first sexual intercourse and last time they had sexual intercourse. These questions were used to improve the accuracy and reliability of the *ever had sex* question.

The mediating variables were measured by multiple item scales that used a 5-point Likert response format, usually ranging from strongly disagree to strongly agree. The scale *abstinence values* (six items) addressed students' commitment to maintaining sexual abstinence until marriage and their acceptance of the idea that marriage is the most appropriate context for sexual activity. *Abstinence efficacy* (six items) assessed students' confidence in their ability to resist pressure to have sex, to avoid situations that would compromise their abstinence position, and to disengage from people who try to pressure them to have sex. *Future impact* (five items) measured students' perception of the impact of sexual involvement on their future education, career, marriage and family life. *Justifications for sex* (five items) measured the rationalizing that adolescents often engage in to legitimize their initiation of sexual activity, such as "being in love" or adopting "safer sex" practices. *Behavioral intentions* (two items) measured students' level of intent and commitment to abstain from sexual activity in the coming year and/or until marriage. Test–retest reliability estimates for these scales ranged from .67 to .75, and internal consistency (Cronbach's alpha) were between .76 and .85.

Analysis

Hypothesis 1. To see how well mediating measures would explain variance in *behavioral intention*, a linear regression model was estimated. The relative contribution of these targeted program variables, versus the group of demographic variables, was determined by partitioning the shared variance based on the proportion of unique variance attributable to each block.

Hypothesis 2. A logistic regression model predicting sexual initiation was estimated. Grade, gender, race, repeated grade, and scores on the behavioral intention measure at baseline were entered into the equation. Categorical variables were effect coded, such that the sum of the category codes within a variable was zero. The intention measure was centered at the mean. These coding revisions aid interpretation of the model by allowing an assessment of the odds-ratio of initiating sexual activity at the predetermined levels of the covariates and given the mean on *behavioral intention.* Program and control respondents who reported they had never had sex as of the pretest were selected and whether or not they initiated sex by the follow-up was predicted. Entering *behavioral intention* measures at the two time points allows an assessment of whether posttest scores, controlling for pretest, are predictive of initiation into sexual activity. This assesses the potency of *behavioral intention* scores (both the level and the change) in predicting *sexual initiation.*

Hypothesis 3. This hypothesis addresses the program's effect on the intervening variables. Separate, repeated-measures ANOVAs were conducted for each mediator variable and each of the three possible time periods (pre to post, pre to follow-up, and post to follow-up). This resulted in a series of 2 X 2 analyses for each time period. These time periods were analyzed separately in order to clarify the pattern of change across time. Given the program and comparison group differences on race and grade, these were controlled for in the repeated measures design. The key result for each of these analyses was the "time by program" effect, which indicates whether the pre–post, the pre–follow-up, or the post–follow-up change was different between the program and the comparison group.

In addition to the change score analysis, a univariate ANOVA was also employed to test the significance of the difference between the program and comparison group scores on the mediator variables at the 12-month follow-up. This tested whether the 12-month follow-up score on the mediating variables was different for the program than for the comparison

group, after controlling for pretest differences on the same mediator scale, with race and grade as covariates.

Finally, we performed hierarchical (multilevel) analysis (Raudenbush and Bryk, 2001) to see if the results of the previous analyses were stable. Specifically, we added random intercepts for school and controlled for the repeated measures among individuals to ensure that observed variation between groups was not better accounted for by the clustering of students within schools or correlations between observations.

Hypothesis 4. A logistic regression analysis was used to test whether program participation was a significant predictor of the likelihood of initiating sex by the 12-month follow-up, for those who were virgins at pretest. The categorical variables of grade and race were entered as covariates. Race was effect-coded into two categories, black and white. Grade was coded using a simple contrast to compare the odds of sexual experience given a grade of 8 or 9 compared to seventh-grade students. The five scales were centered at the mean and entered as covariates in order to control for the amount of the initiation-rate differences attributable to the pretest differences on these measures. The program variable was coded as a zero for the comparison group and as a one for the program group. These coding revisions aid interpretation of the model. Specifically, the odds-ratio associated with the intercept value indicates the odds that a comparison group individual will initiate sex, given an average on all of the other covariates. These odds can then be compared to the odds that the average program student will initiate sex, given an average on all other covariates.

Program improvement process. The final focus of the evaluation was to assess the Heritage Keepers'® self-correcting mechanisms for maintaining a high caliber program. We reviewed two aspects of the program's system. First, the data that is collected about individual teachers and how it is used to improve overall teaching effectiveness, and, secondly, the monitoring and feedback that occurs with regard to fidelity-to-plan in field settings.

Results

Hypothesis 1

The analysis reported in Table 3 indicates that 60.7 percent of the variance in behavioral intention was explained by all of the variables. We calculated that most of the variance, 49.4 percent, was uniquely explained by the mediating variables, about 1.6

percent was uniquely explained by the demographic variables, and the remaining variance was shared between the two blocks of variables. The shared variance, when proportionately distributed, brings the variance explained by the mediating variables to 58.8 percent. The demographic variable, with its proportion of shared variance added, brings its contribution to 1.9 percent. The model with the mediators entered first is reported in table 2.

Hypothesis 2

Results of the logistic regression analysis showed a significant relationship between virgin students' *behavioral intention* scores and their initiation of sexual intercourse one year later. Ignoring program involvement and controlling for demographic covariates, the estimated odds of initiation of sexual intercourse after one year were .208 (table 3). The odds of initiation drop to .115 when *behavioral intention* scores one unit above the mean is factored in (.552 x .208 = .115) indicating that students scoring higher on the *behavioral intention* measure were significantly less likely to initiate sex after one year.

Hypothesis 3

Table 4 contains the means and standard deviations of the pre, post, and 12-month follow-up scores on all five mediating variables for the program and comparison group. It also contains the F-values and significance levels for the time-by-program analysis for each of the three time periods. For the first time period (pre to post), the time-by-program interaction term was significant for all five mediating variables. The effect sizes (Cohen's *d*) for the mediating variables in this time period were calculated at .42 for *abstinence values*, .34 for *behavioral intentions*, .36 for *future impact*, .44 for *justifications for sex*, and .35 for *abstinence efficacy*.

For the second time-period analysis (pre to follow-up), the program-by-time interaction term was significant for three of the five mediating variables, and in the expected direction. This effect appears to be the result of a pattern whereby the program students made a significant pre–post gain over the comparison students and maintained that difference over time. Two mediating variables where this pattern did not occur were *justification for sex* and *abstinence efficacy*. The pattern for these variables did not show the same decline over time (from post to follow-up) among the control group as was found for *behavioral intention, abstinence values,* and *future impact*.

The positive pre–post change achieved on the mediating variables was not expected to last indefinitely. Deterioration from post to follow-up (our third time-period analysis) was found on all of the mediators for the program group and for three of the mediator-scale scores in the comparison group. The amount of deterioration in scores over 12 months was similar for both groups for three of the mediating variables (*abstinence values, future impact,* and *behavioral intentions*) scores. Most importantly, the differences between groups in scores on these variables at posttest were still apparent at the follow-up 12 months, with *justification for sex* and *abstinence efficacy* being the exceptions (last column, table 4).

When we ran all of the same comparisons as above using the multilevel models, we found essentially the same patterns. While significant variance was accounted for by the school and within subject correlations, the sign and significance of the program to comparison group tests were the same in every case except for one. In the repeated-measures designs, *justifications for sex* did not show a significant effect from pre to follow-up. In the multilevel model, it was marginally significant ($b = .065, p = .067$).

Hypothesis 4

Results of this logistic-regression prediction of sexual initiation rates (table 5) indicated that the odds of initiation for the average comparison student were .271. The odds of initiation for the average program participant were .146 (.539 * .271 = .146), significantly lower than the comparison group odds ($p < .001$). More concretely, of 1,216 program students who were virgin at the pretest and who also answered the follow-up sex question, 176, or 14.5 percent, had sex between the pre and the follow-up. Of 253 virgin comparison students, 67, or 26.5 percent, initiated sex between the pre and the follow-up. The model containing the program variable and covariates was compared to the model containing only the covariates. The difference in the model chi-squares was statistically significant ($\Delta Chi^2 = 11.2$ with $df = 1, p < .001$), indicating that the program variable explains statistically significant variance in initiation rates.

Program-improvement process

Survey data recording students' pre–post change on key mediating variables is tied to the individual Heritage program teachers. Statistical analysis is performed at the teacher level, with tests for significant pre–post differences by teacher. Each teacher is given feedback at mid-year on their students'

outcomes (table 6) so they can make mid-year adjustments in teaching methods where needed. Teachers with strong results on the mediating-variable concepts were asked to present some of their methods in meetings with all teachers. In addition, teachers that demonstrated weak improvement on specific scales were paired with, and trained by, a teacher who demonstrated strong improvement on that scale. Teachers were required to develop a personalized improvement plan, approved by the central office, and which they were expected to follow.

The program's fidelity to plan is monitored frequently through means of a telephone check-in with teachers by the Heritage Keepers® central staff every two weeks. Teachers report progress on their personalized teaching improvement plans, including the content of lessons and fidelity to the program, as well as teaching methods. Teachers are given feedback immediately if a pattern of deviation appears. Starting with school year 2005–2006, this information will be recorded on a daily log that indicates which students were in attendance and what topics were covered. These logs will be entered in a data file, allowing for a more systematic analysis and use.

Discussion

The results of this study provide support for all four hypotheses. Most notably, program virgins were about one-half as likely (odds ratio=.539) as comparison group virgins to initiate sex by the 12-month follow-up, after controlling for pretest differences (on race, grade, and all mediating variables). The results also demonstrated the importance of several mediating variables that explain the way in which the program produced the outcome of delayed initiation. Specifically, we found that a self-report measure of *behavioral intention* was predictive of initiation of sexual activity. *Behavioral intention,* in turn, was strongly related to other constructs central to the behavioral influence model on which the intervention was based: *abstinence values, justification for sex, abstinence efficacy,* and perceived *future impact of sex.*

Further analyses showed that the change in these mediating constructs was consistent with the difference between groups in sexual initiation rates at 12-month follow-up. As might be expected, the short-term results dissipated over the 12-month follow-up absent ongoing support and reinforcement. The dissipation rate was similar for program and comparison groups. However, the between group differences in three of the five mediating variables observed at post test were maintained for one year, and those differences corresponded to differences in

sexual initiation rates measured at the 12-month follow-up. This lends further support to the explanation that the program, through direct impact on targeted mediating factors, was responsible for the observed initiation rate differences between treatment and comparison groups. The results of the multilevel models analysis increased confidence that the observed improvements on mediator variables were reflective of a program effect and not a failure of the analyses to appropriately estimate the model. In addition, the program effects on the mediator scores were equally strong among males and females, blacks and whites, and seventh, eighth, or ninth graders.

An obvious alternative explanation for initiation rate differences is that they were related to sample differences between the two groups. It is important to determine whether the treatment and comparison groups were affected by either selection bias or attrition bias so that the observed initiation rate differences between the two groups can more confidently be attributed to program effects rather than sample differences, and if the treatment and comparison groups are different on relevant variables, to account for or control for those differences.

Attrition bias could affect group comparability if it results in significant differences on the key measures. In this sample, the attrition that occurred did not increase the dissimilarity between the treatment and comparison groups, but did leave both groups in the final sample somewhat less risk prone than the original sample. The end result is that the program effect is not undermined by attrition, but there is a limitation in generalizing the findings to the original sample.

This leads to a further examination of starting differences or selection bias as a contributing factor in program outcomes. At pretest, both the program and comparison group had a higher proportion of females than males, and the comparison group had a higher proportion of minorities, more ninth graders, and lower scores on the mediating variables than the program group (table 1). Since the analysis examining initiation rates for virgin students controlled for grade, race, and mediator-variable scores at pretest, the significant program effect reported has already accounted for the preexisting differences between groups.

As a further test of the impact of sample differences, race and gender groups were examined separately (grade level was not, since it was not a significant predictor of initiation *rates*). The difference between program and comparison students' initiation rates was large among both blacks and whites. Specifically, 30.6 percent of blacks in the comparison group initiated sex by the follow-up compared to 17.4 percent of

blacks in the program (chi square =12.6, $p<.001$). The numbers for whites were 20.8 percent of comparison students versus 12.0 percent of program students (chi square=6.03, $p=.014$). Looking at odds ratios separately by race produced a similar result. The odds ratio of initiation for program blacks versus comparison blacks was .498, compared to an odds ratio for whites of .472, after controlling for pretest differences in risk level. Thus, it does not appear that the overall group differences found in initiation rates were accounted for by the lower proportion of blacks in the program group. The initiation rate differences between the program and comparison group were also statistically significant for both boys (19.7 percent vs. 31.9 percent, $p<.05$) and girls (11.3 percent vs. 23.5 percent, $p<.05$), suggesting that the program effect was not related to the higher proportion of girls in the sample.

A final test of sample effects was conducted by creating a subset of the program sample that attempted to correct for the pretest differences that existed between groups in the full sample. First, a subset was created that matched on demographic variables (gender, race, grade, and repeated grade). Then within each of those cells (seventh-grade white males, etc.) we produced exact matches on the five mediating variables at baseline, using propensity scores obtained from a logistic regression to predict treatment or comparison group membership. This produced program and comparison groups that were extremely well matched on variables predictive of the outcome measure of sexual initiation (no significant differences). Examination of the initiation rates at one-year follow-up revealed a 13.9 percent rate for the program group, and 25.6 percent for the comparison group (odds ratio=.444, $p<.001$) results that were nearly identical to (or slightly better than) those for the full, but less well matched, sample.

After controlling for these pretest differences in the demographic and mediating variables using three different methods, the program students were still significantly less likely than comparison students to initiate sex at the 12-month follow-up, suggesting that the initiation rate differences were not attributable to the pretest differences of the sample.

Finally, the Heritage program takes seriously the implementation and process issues as factors in program success. It has developed a fairly rigorous system of tracking the level of implementation in the field, measuring the quality of teaching, and providing regular data-based feedback to the appropriate staff in productive settings.

Limitations

It is possible that, in spite of best efforts to control for differences between program and comparison groups, these groups differ in some ways not measured by our study, and that these differences are contributing to the reduction in initiation rates. This also limits our ability to measure the true effect size of program impact. In addition, the high attrition rate limits our ability to generalize the findings to a higher-risk population. These limits on the interpretation of results may be addressed with better-matched comparison groups or a random assignment design, and with better long-term follow-up of individuals.

Program implications and recommendations

Testing the connection between key mediator variables and the transition into risky behavior helps to establish a better causal model for adolescent risk avoidance. The significant drop in mediator scores from post to follow-up has important program delivery implications in terms of maintenance and reinforcement. This argues for strategies that will extend, support, and reinforce the intervention's impact over time—no easy task in a school setting where multiple demands are competing for student time and attention.

Conclusions

The results of this study suggest that a carefully developed abstinence-centered education program can lower the rate at which virgin youth initiate sex. The Heritage program produced a significant and substantial delay in sexual initiation 12 months after the intervention. In addition, a better understanding of the mechanisms that produced this change was realized by including hypothesized mediating factors in the program design and evaluation. These findings support the premise that primary prevention (risk avoidance) efforts to influence teens towards sexual abstinence are a viable strategy.

References

Ajzen, I. 1991. The theory of planned behavior. *Organizational Behavior and Human Decision Processes 50:* 179–211.

Armitage, C., M. Conner. 2000. Social cognition models and health behavior: A structured review. *Psychology and Health* 15: 173–89.

Bandura, A. 2004. Health promotion by social cognitive means. *Health Education & Behavior* 31: 143–64.

Centers for Disease Control and Prevention. 2001. Tracking the hidden epidemics 2000: Trends in STDs in the United States. U.S. Department of Health and Human Services, Atlanta, GA. www.cdc.gov/nchstp/od/news/RevBrochure1pdfintro.htm. Accessed March 2003.

Centers for Disease Control and Prevention. 2003. Sexually transmitted disease surveillance, 2002. U.S. Department of Health and Human Services, Atlanta, GA. www.cdc.gov/std/stats/tables/table12B.htm. Accessed February 2004.

Conner, M. and C. J. Armitage. 1998. Extending the theory of planned behavior: A review and avenues for further research. *Journal of Applied Social Psychology* 28: 1429–64.

Floyd, D. L., S. Prentice-Dunn, and R. W. Rogers. 2000. A meta-analysis of research on protection motivation theory. *Journal of Applied Social Psychology* 30: 407–29.

Genuis, S. J., and S. K. Genuis. 2004. Managing the sexually transmitted disease pandemic: A time for reevaluation. *American Journal of Obstetrics and Gynecology* 191: 1103–12.

Godin, G., and G. Kok. 1996. The theory of planned behavior: A review of its application to health-related behaviors. *American Journal of Health Promotion* 11: 87–89

Grunbaum, J., L. Kann, S. Kinchen, J. Ross, J. Hawkins, and R. Lowry. 2004. Youth risk behavior surveillance–United States, 2003. *MMWR Surveillance Summaries.* Centers for Disease Control and Prevention, U.S. Department of Health and Human Services, Atlanta, GA. www.cdc.gov/mmwr/preview/mmwrhtml/ss5302a1.htm. Accessed September 2004.

Hall, D. 2005. Getting honest about grad rates: How states play the numbers and students lose. Education Trust, Washington, DC. www2.edtrust.org. Accessed July 2005.

Jaffee, S. R. 2002. Pathways to adversity in young adulthood among early child bearers. [Electronic version] [Abstract]. *Journal of Family Psychology* 16: 38–49.

Kirby, D. 2001. Emerging answers: Research findings on programs to reduce teen pregnancy. National Campaign to Prevent Teen Pregnancy, Washington, DC.

———. 2002a. Antecedents of adolescent initiation of sex, contraceptive use, and pregnancy. *American Journal of Health Behavior* 26: 473–85.

———. 2002b. Do abstinence-only programs delay the initiation of sex among young people and reduce teen pregnancy? National Campaign to Prevent Teen Pregnancy, Washington, DC.

Manlove, J. M., E. Terry-Humen, A. Papillo, K. Franzetta, S. Williams, and S. Ryan. 2002. Preventing teenage pregnancy, childbearing, and sexually transmitted diseases: What the research shows. In *American Teens: A Special Look at "What Works" in Adolescent Development,* ed. Child Trends and John S. and James L. Knight Foundation (pp. 6–23). Washington, DC: Child Trends.

Maynard, R. A. (ed.). 1997. *Kids Having Kids: Economic Costs and Social Consequences of Teen Pregnancy.* Washington, DC: Urban Institute.

Miller, B. C., and K. A. Moore. 1990. Adolescent sexual behavior, pregnancy, and parenting: Research through the 1980s. *Journal of Marriage and the Family* 52: 1025–44.

National Campaign to Prevent Teen Pregnancy. 2003. America's adults and teens sound off about teen pregnancy: An annual national survey. Washington, DC. www. teenpregnancy.org. Accessed January 2004.

National Institute of Allergy and Infectious Diseases. 2001. Workshop summary: Scientific evidence on condom effectiveness for sexually transmitted disease (STD) prevention. National Institutes of Health, Department of Health and Human Services, Washington, DC. July.

Plotnik, R. D. 1992. The effect of attitudes on teenage premarital pregnancy and its resolution. *American Sociological Review* 57: 800–11.

Raudenbush, S. W., and A. S. Bryk. 2001. *Hierarchical Linear Models: Applications and Data Analysis Methods.* Advance Quantitative Techniques in the Social Sciences Series. Thousand Oaks, CA: Sage.

Rector, R. 2000. The effectiveness of abstinence education programs in reducing sexual activity among youth. Backgrounder 1533, Heritage Foundation, Washington, DC. http://www.heritage.org/Research/Family/BG1533.cf m. Accessed August 2003.

Resnick, M. D., P. S. Bearman, R. W. Blum, K. E. Bauman, K. M. Harris, and J. Jones. 1997. Protecting adolescents from harm: Findings from the National Longitudinal Study on Adolescent Health. *Journal of the American Medical Association* 278: 823–32.

Scher, L. S., R. A. Maynard, and M. Stagner. 2005. Interventions intended to reduce pregnancy-related outcomes among adolescents. Unpublished manuscript. June.

Strecher, V. J., B. M. DeVellis, M. H. Becker, and I. M. Rosenstock. 1986. The role of self-efficacy in achieving health behavior change. *Health Education Quarterly* 13: 73–92.

Sulack, P. J. 2003. Sexually transmitted diseases. *Seminars in Reproductive Medicine* 21: 399–413.

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

U.S. Department of Health and Human Services. 2000. *Healthy People 2010* (2nd ed.), with *Understanding and Improving Health and Objectives for Improving Health.* 2 vols. Washington, DC: U.S. Government Printing Office. November.

Weed, S. E., and J. A. Olsen. 1988. Policy and program considerations for teenage pregnancy prevention: A summary for policy makers. Family Perspective, 22, 235–252.

Weed, S. E., J. A. Olsen, J. De Gaston, and J. Prigmore. 1992. Predicting and changing teen sexual activity rates: A comparison of three Title XX programs. Office of Adolescent Pregnancy Programs, Department of Health and Human Services, Washington, DC.

Weinstock, H., S. Berman, and W. Cates. 2004. Sexually transmitted diseases among American youth: Incidence and prevalence estimates, 2000. *Perspectives on Sexual and Reproductive Health 36*: 6–10.

Table 1. Demographic distributions of program and comparison groups

Pretest data for virgins only, linked pre-to-follow-up

Variable	Program N=1,281 (percent)	Comparison N=254 (percent)	Test of statistical significance
Grade			
7	664 (53.2)	85 (33.5)	3.42***
8	238 (19.1)	48 (18.9)	.03
9	346 (27.7)	121 (47.6)	4.0****
Gender			
Female	776 (62.2)	162 (63.8)	0.38
Male	472 (37.8)	92 (36.2)	0.29
Race			
Black	569 (45.6)	147 (57.9)	2.66**
White	679 (54.4)	107 (42.1)	2.37*
Repeated grade			
Yes	204 (16.4)	47 (18.7)	0.38
No	1038 (83.6)	205 (81.3)	0.81
Abstinence efficacy	3.74	3.83	1.7
Abstinence values	3.82	3.60	11.6***
Future impact	3.54	3.21	22.3****
Justification for sex	3.78	3.59	10.1***
Behavioral intentions	3.73	3.50	9.4**

Note: Demographic variables tested using Hintze's test of the difference in two independent proportions. Mediating variables tested with univariate F ratio.

* $p<.05$, ** $p<.01$, *** $p<.001$, **** $p<.0001$

Table 2. Linear regression analysis explaining variance in pretest behavioral intentions

N = 1,464 list-wise cases, virgins only

Source	B	SE B	β
Constant	0.60	0.15	
Step 1			
Values	0.37	0.03	0.32***
Efficacy	0.28	0.02	0.26***
Impact	0.07	0.02	0.07**
Justification	0.34	0.03	0.28***
Step 2			
Grade	−0.11	0.02	−0.09***
Race	−0.19	0.04	−0.09***
Gender	−0.15	0.04	−0.07***
Repeated Grade	−0.01	0.05	−0.00

Note. *adj. R^2* = 0.592 for Step1; Δ *adj. R^2* = 0.016 for Step 2 ($p < .001$). * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 3. Logistic regression analysis predicting initiation of sex at the follow-up

N=1,291, virgins only

Source	df	B	SE	Exp(B)
Grade				
8 vs. 7	1	−0.06	0.14	0.94
9 vs. 7	1	0.08	0.12	1.08
Race	1	0.38	0.17	1.46*
Gender	1	0.28	0.17	1.32
Repeated grade	1	0.79	0.19	2.19**
Centered pretest behavioral intention	1	−0.09	0.10	0.92
Centered post test behavioral intention	1	−0.60	0.10	0.552**
Constant	1	−1.57	0.11	0.208**

Note. * $p < .05$, ** $p < .01$.

Table 4. Mean scores on intermediate variables and results of repeated-measures analyses of variance

Virgins only		Mean (standard deviation)			F of time by program effect			F
Scale	Group	Pre	Post	Follow-up	Pre–post	Pre–follow-up	Post–follow-up	Follow-up
Abstinence values	Program	3.81 (.88)	4.08 (.83)	3.84 (.92)	37.1***	32.2***	.04 NS	64.6***
	Comparison	3.61 (1.10)	3.49 (1.08)	3.27 (1.20)				
Behavioral intention	Program	3.74 (1.04)	3.98 (1.03)	3.65 (1.12)	25.1***	12.8***	.40 NS	24.8***
	Comparison	3.50 (1.18)	3.43 (1.19)	3.16 (1.22)				
Future impact	Program	3.54 (1.00)	3.90 (.95)	3.61 (1.01)	26.9***	31.5***	.17 NS	67.0***
	Comparison	3.22 (1.08)	3.22 (1.13)	2.92 (1.21)				
Justification for sex	Program	3.77 (.87)	4.06 (.82)	3.82 (.85)	40.4***	.66 NS	40.1***	.72 NS
	Comparison	3.60 (.91)	3.51 (.95)	3.70 (.97)				
Abstinence efficacy	Program	3.75 (1.00)	4.00 (.92)	3.77 (1.02)	25.6***	3.3 NS	8.9**	2.4 NS
	Comparison	3.81 (.96)	3.74 (1.04)	3.75 (.98)				

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

The last column represents the F value and significance level for between-group difference at follow-up, controlling for grade, race, and the pretest score.

Table 5. Logistic regression prediction of sexual initiation rates, controlling for pretest scores, grade, and race

N=1,448, virgins only

Source	df	B	SE	Exp(B)
Grade				
8 vs. 7	1	−.105	.207	0.90
9 vs. 7	1	.085	.176	1.09
Race				
Black vs. White	1	.427	.155	1.53**
Centered pretest predictor scores				
Abstinence values	1	.017	.124	1.02
Behavioral intention	1	−.147	.108	0.86
Future impact	1	−.09	.095	0.91
Justification for sex	1	−.301	.119	0.74**
Abstinence efficacy	1	−.231	.094	0.79**
Program participation	1	−.617	.181	0.539**
Constant	1	−1.61	.098	0.271***

Note. $* p < .05$, $** p < .01$, $*** p < .001$.

Table 6. Student outcomes on mediating variables by individual teacher, fall 2004–05

	Strong Teacher (student N=152) PRE POST F-value Sign.	Weaker Teacher (student N=48) PRE POST F-value Sign.
Abstinence values	3.81 – 4.44	2.98 – 3.11
	97.89 ****	3.15 (*)
Abstinence Efficacy	4.10 – 4.58	2.79 – 3.04
	53.28 ****	1.21 NS
Future impact	3.52 – 4.23	2.72 – 2.91
	121.09 ****	5.23 *
Independence from peers	4.05 – 4.32	3.47 – 3.35
	24.01 ****	2.79 NS
Justification for sex	3.91 – 4.48	2.91 – 3.17
	73.03 ****	5.86 *
Behavioral intention	4.03 – 4.60	2.65 – 2.61
	80.95 ****	.04 NS

(*) $p<.10$, * $p<.05$, **** $p<.0001$, NS = not significant