Effectiveness of Prevention Programs for Adolescent Pregnancy: A Meta-Analysis

Cynthia Franklin; Darlene Grant; Jacqueline Corcoran; Pamela O'Dell Miller; Linda Bultman

Journal of Marriage and the Family, Vol. 59, No. 3. (Aug., 1997), pp. 551-567.

Stable URL:
http://links.jstor.org/sici?sici=0022-2445%28199708%2959%3A3%3C551%3AEOPPFA%3E2.0.CO%3B2-3

Journal of Marriage and the Family is currently published by National Council on Family Relations.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/NCFR.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org
Fri Jul 13 13:44:50 2007

CYNTHIA FRANKLIN AND DARLENE GRANT
University of Texas at Austin

JACQUELINE CORCORAN *University of Texas at Arlington**

PAMELA O'DELL MILLER *University of Texas at Austin***

LINDA BULTMAN *Texas Department of Health****

Effectiveness of Prevention Programs for Adolescent Pregnancy: A Meta-Analysis

Using meta-analysis, we analyzed 32 outcome studies on the primary prevention of adolescent pregnancy and examined several moderator variables in relationship to the findings. Three outcome variables—sexual activity, contraceptive use, and pregnancy rates or childbirths—were analyzed as three separate and independent meta-analyses. Results indicate that the pregnancy prevention programs that we examined have no effect on the sexual activity of adolescents. We found sufficient evidence to support the efficacy of pregnancy prevention programs for increasing use of contraceptives. A smaller but significant amount of evidence supports program effectiveness in reducing pregnancy rates.

School of Social Work, University of Texas at Austin, 1925 San Jacinto Boulevard, Austin, TX 78712 (cfranklin@mail.utexas.edu).

*School of Social Work, University of Texas at Arlington, 211 South Cooper, Arlington, TX 76019.

**2949 Waterloo Road, Pearland, TX 77581.

***Texas Department of Health, 1100 West 49th Street, Austin, TX 78756.

Key Words: adolescent pregnancy prevention, meta-analysis.

Political debates rage about the effectiveness of pregnancy prevention programs for adolescents and whether such programs should be offered (Whitehead, 1994). Several narrative or qualitative reviews have examined the findings of pregnancy prevention studies. Narrative reviews, however, have been limited by their lack of comprehensiveness in study selection (Jorgensen, Potts, & Camp, 1993) or by simply citing results of individual programs without attempting to synthesize the literature in a systematic way (Beck & Davies, 1987; Hofferth, 1991). They fail to provide systematic mechanisms for combining the results of different studies and for weighting and assessing the impact of variations in study methodology. Another failure of narrative reviews is that they are not equipped to assess the variation within studies of programs' effects on different subpopulations, such as specific gender, ethnic, and age groups. In addition, narrative reviews frequently suffer from mixed, contradictory, and inconclusive results. Meta-analysis may address these limitations because it allows studies to be systematically quantified and their results summarized as a common effect. This article reports the findings from a meta-analysis that evaluated the effectiveness of pregnancy prevention programs. We summarize narrative reviews conducted in this area, and we

examine the potential of a meta-analysis to address the limitations of these narrative reviews. Finally, we present the methodology and the results of the meta-analysis, itself.

Narrative Reviews

The most recent comprehensive narrative review of the literature involves school-based programs to prevent adolescent pregnancy (23 studies included) conducted by Kirby and colleagues (1994). The review narratively synthesizes research on the effectiveness of programs, identifies distinguishing characteristics of effective programs, and points out important research questions in this area. To be included in the review, a study had to be accepted for publication or published in a peer-reviewed journal. Findings, however, are inconsistent and inconclusive. For example, Kirby et al. identify three published studies on abstinence-based programs and conclude that the evidence is mixed whether they affect the initiation of sexual activity or contraceptive behavior. The national survey data examined also demonstrated an inconsistent relationship between sex education and sexual activity. The authors conclude that the mixed results from research studies make it impossible to "determine the impact of school-based or school-linked reproductive health services" (p. 356).

In a prior evaluation of school-based clinics, Kirby (1992) concluded that the reduction in birth rates initially reported for clinics in St. Paul, Minnesota, was not borne out by subsequent analyses. In a study of six school-based clinics, Kirby, Waszak, and Ziegler (1991) report that ". . . simply dispensing contraceptives was not sufficient to dramatically increase contraceptive use, but more comprehensive approaches that focus on pregnancy and AIDS prevention might have a greater impact" (p. 284).

Dawson (1986) concludes from a review of the pregnancy prevention literature that "the existing data do not yet constitute consistent, compelling evidence that sex education programs are effective in increasing teenage contraceptive use and reducing adolescent pregnancy" (p. 163). Dawson indicates that the most successful programs for influencing contraceptive behavior are based in a clinic. Additionally, there is some evidence that clinic-based programs can play a role in the reduction of pregnancy rates, although this finding may be confounded by adolescents who receive abortions instead of learning how to use contraceptives.

Hofferth (1991) concurs that community-based and school-linked clinics can reduce pregnancy rates. She states, however, that the same cannot be said for school-based clinics. Although school-based clinics appear to improve the general level of health care to students, no significant effects on sexual activity, contraceptive use, or pregnancy rates have been indicated consistently. From Hofferth's review, it may be inferred that community-based programs are more effective than school-based clinics for changing specific behavioral outcomes in pregnancy prevention.

Beck and Davies (1987) reviewed 12 community and school-based programs that affect contraceptive use. They list findings of each study and conclude that those targeting communication and problem-solving skills are the most effective. Howard and McCabe (1992) also tout programs that focus on building skills to resist social and peer pressure to engage in sexual intercourse. These programs are more effective than programs that simply educate adolescents about the risks and consequences of sexual behavior. Postrado and Nicholson (1992) and Kirby, Barth, Leland, and Fetro (1991) claim that, in at least some studies, the early promise of skills training based on social-cognitive theory seems to be supported by subsequent empirical research.

The narrative reviews have yielded mixed, contradictory, and inconclusive results concerning the outcomes of programs to prevent teen pregnancy. Furthermore, these qualitative reviews have difficulty synthesizing the results of studies that differ in essential program components—such as type of program, site of service delivery, and intended participant. A meta-analysis can address these limitations.

Meta-Analysis

Meta-analysis allows the outcomes of various studies to be summarized statistically on a common scale of effect size (Glass, 1976). An effect size is a type of standard score that permits the combining of results from studies and is usually determined by dividing the difference between an experimental group and a control group by means the standard deviation of the control group. Summarizing studies in this way reveals which interventions demonstrate the most effects.

Although meta-analysis is not without its limitations, it has several advantages, compared with

narrative reviews. In contrast to Frost and Forrest (1995) who argue that it is premature to conduct a comprehensive meta-analysis, we believe that a meta-analysis is overdue. Narrative reviews in this area rely on what has been referred to as the "vote-counting method" (Hunter & Schmidt, 1990; Light & Smith, 1971). In the vote-counting method, studies are categorized as positive significant, negative significant, or nonsignificant; the modal category determines the best representation of outcome. The major limitation of this method is that it can reliably discover a treatment's effect size only if the studies involved have equal sample sizes and a unimodal distribution indicative of one population, a relatively rare phenomenon. More commonly, sample sizes of studies differ and are derived from multimodal distributions, which presents problems for narrative reviews. Furthermore, statistical significance is a function of both treatment effect and sample size. If the sample sizes of studies differ, findings from larger samples are more likely to reach significance even when the actual treatment effect is small or equal to those in other studies (Bangert-Drowns, 1986).

Further, if study findings are nonsignificant but still in the direction of the hypothesis, they are viewed as failures to confirm the hypothesis without considering that the statistical power of the studies may be weak. In such cases, extremely large effects would have to be achieved in order to reach significance, given the sample size (Cook & Leviton, 1980). In addition, the insensitivity to interactional effects is a crucial limitation to traditional narrative reviews. Meta-analysis is not hampered by such problems because it uses summary statistics (e.g., proportions, means, or correlations) of the studies' dependent measures as the unit of analysis, rather than relying on frequency counts or statistical significance within studies. Aggregate data are then analyzed quantitatively (Cook & Leviton, 1980). Finally, another advantage of meta-analysis is that the impact of moderator variables on the independent variables (Bangert-Drowns, 1986) may be investigated.

We found only three quantitative reviews or meta-analyses of the literature on the prevention of adolescent pregnancy. The most recent review conducted by Frost and Forrest (1995) was not published until after we had completed our meta-analysis. Frost and Forrest utilize some of the same studies used in our meta-analysis, and they assess similar outcome measures. Frost and Forrest specifically address the efficacy of teen pregnancy prevention programs in relationship to their impact on sexual activity, contraceptive use, and pregnancy or childbirth rates. They summarize the nonstandardized effect sizes across five studies of school-based pregnancy prevention programs. Only five studies were chosen because Frost and Forrest believed they represented rigorously evaluated programs. The authors conclude that "all four of the programs that measured adolescent sexual and contraceptive behavior delayed the initiation of sexual activity among many teenagers, and three of these four programs increased the proportion of sexually active teens using contraceptives" (pp. 194–195). However, their results do not address the programs' effectiveness on pregnancy and childbirth rates.

Frost and Forrest's study represents a first attempt at statistically synthesizing the outcomes of programs to prevent adolescent pregnancy. The limited number of studies reviewed, however, means that it does not represent a comprehensive synthesis of existing research literature. And Frost and Forrest do not follow statistical procedures commonly associated with a meta-analysis. The usual procedure is to include as many studies as possible from a substantive area. Variables that may contribute to treatment effects are systematically coded. Spurious effects such as design quality also are examined. Investigating the impact of moderator variables—such as age, gender, and ethnicity—is also an important contribution of a meta-analysis. The Frost and Forrest study is limited by statistical methods that do not allow standardized comparisons, weighting of sample sizes, the examination of moderator variables, and other critical analysis commonly associated with meta-analysis.

Iverson and Levy (1982) published a quantitative review that was concerned with a synthesis of a group of adolescent pregnancy education programs from 1970 to 1980. The purpose of their review is to illustrate the use of a meta-analysis in health education research. This is not an outcome review. It mainly serves as an example of how to do meta-analysis. Fourteen studies are reviewed and grouped into six outcome variables. Adolescent pregnancy prevention yields small to moderate effect sizes on the six outcomes.

Whitley and Schofield (1986) conducted a meta-analysis on 134 studies of adolescent contraceptive use among both high school and college adolescents. College women are the most frequently represented group (57%). The studies relate to two explanatory models, a career model and a decision model, which the authors found

complimentary. Both theoretical models found support in the meta-analysis. Although quantitative reviews of research on teen pregnancy outcomes are sparse, the three meta-analyses that have been conducted offer more consistent support for critical variables associated with the prevention of teen pregnancy and its outcomes than the qualitative studies.

Our meta-analysis represents a comprehensive effort to evaluate the effectiveness of primary prevention programs aimed at junior high and high school teens. This meta-analysis summarizes the findings of 32 studies and uses the outcome variables of sexual activity, contraceptive use, and rates of pregnancy or childbirths.

Methodology

A search of the literature in family sciences (Social Sciences Index), psychology (PsychLIT), education (Educational Resources Information Center), medicine (MEDLINE), and nursing (Cumulative Index to Nursing and Allied Health Literature) was undertaken in order to identify outcome studies on primary pregnancy prevention programs with adolescents. To assure retrieval of the highest quality studies, we limited our search to peer-reviewed, refereed journal articles in the published literature. Our only deviation from this rule is the inclusion of five studies that appeared in a 1992 publication reviewing effective studies of adolescent pregnancy prevention (Eisen & Zellman, 1992; Howard & McCabe, 1992; Nicholson & Postrado, 1992; Philliber & Allen, 1992; Thomas et al., 1992). Our justification was that these studies represented and were reviewed by some of the top researchers in the field.

Initial searches uncovered more than 500 references related to adolescent pregnancy prevention published through 1995. These were screened according to the following criteria: Did they involve studies of program outcomes? A majority of references were dropped because they did not meet this criterion. Was outcome assessed and presented behaviorally by either sexual activity, contraceptive use, or pregnancy? Many studies were eliminated from the meta-analysis because they assessed attitudes, knowledge, or self-esteem, rather than behavior. Did studies provide adequate statistical data needed for a meta-analysis? We dropped several studies because they did not meet this criterion. When this screening process was applied, a total of 32 outcome studies in 11 journals and one 1992 publication were located. The Appendix presents a review of these studies and their characteristics, as well as other information pertinent to the results of the meta-analysis.

Because the reports on some studies appeared in different journals or publications (e.g., Zabin, 1992; Zabin, Hardy, Streett, & King, 1984; Zabin, Hirsch, Smith, Streett, & Hardy, 1986a, 1986b), we used only unreplicated data for the meta-analysis. Although every effort was made to be exhaustive in our literature search, it is possible that the studies in the meta-analysis are not completely representative of the field.

Criteria for Inclusion in the Meta-Analysis

All studies in the meta-analysis have adolescents as their subjects. Adolescents were broadly defined to include youths between the ages of 11 and 20 years. All studies have a focus on the primary prevention of adolescent pregnancy. Primary prevention is targeted at adolescent girls who have not previously been pregnant. We used only studies that determined empirically the impact of a particular intervention on one of three behavioral outcome variables—sexual activity, contraceptive use, or pregnancy rates. The rationale for the focus on these outcomes is that previous studies and reviews of the literature found that changes in knowledge and attitudes are not strongly associated with changes in adolescent sexual behavior, which is the primary concern of most prevention efforts (Kirby et al., 1994; Zelnik & Kim, 1982). Although we acknowledge the limitations of these outcome measures (such as the reliability of self-report), we believe that these measures provide a direct assessment of the impact of programs.

Deviations in the measurement of outcomes appear to represent the state of the art in the field. For this reason, we accepted all measures associated with the behavioral dimensions of sexual activity, contraceptive use, and pregnancy rates. Contraceptive use was defined to include any measure or indicator that assessed an adolescent's behavior concerning contraceptive use or practices during sexual intercourse. Contraceptive use outcomes particularly showed a range of definitions, from "effective contraceptive use during the past week" to "ever used contraception." Questionnaires asking adolescents to self-report retrospectively on their use of different types of contraceptives were frequently used.

Sexual activity is defined broadly to include any measure or indicator that assessed an adoles-

cent's frequency of, delay of, or engagement in sexual intercourse. This usually involved questionnaires that asked adolescents to report their sexual activities retrospectively. Pregnancy rates were defined to mean any measure or indicator that assessed a young woman's conceiving or giving birth to a child, or a young man's impregnating a young woman or fathering a child. Definitions varied across studies and included measures such as clinic or hospital records, birth rates, and self-report questionnaires.

Studies also must have included statistical data appropriate to the meta-analysis or at least sufficient numerical information to calculate the appropriate statistics. Because of lack of clarity or missing data, it was necessary for us to use only portions of the data in some studies. For example, the Eisen and Zellman (1992) study was set up as a quasi-experimental design with treatment and comparison groups, pretest, posttest, and follow-up, but due to the difficulty of tracking sample sizes throughout the study presentation, the design had to be converted to posttest only to give us the statistical information necessary for the meta-analysis. We encountered similar problems in deciphering the data in four other studies (Furstenberg, Herceg-Baron, Shea, & Webb, 1984; Nicholson & Postrado, 1992; Shea, Herceg-Baron & Furstenberg, 1984; Zabin et al., 1986a).

Coding the Studies

Two reviewers collected data from the studies. In addition to statistical data, the following moderator variables were coded: (a) locus of intervention—community vs. school-based or school-linked; (b) type of program—clinic vs. nonclinic; (c) type of intervention—abstinence only, sex education with no contraception, sex education with contraceptive knowledge-building and distribution; (d) focus of intervention—building skills or no skills building; (e) design—experimental, quasi-experimental, pretest and posttest, posttest only, and other; and (f) subject-related moderator variables, such as age, gender, ethnicity, socioeconomic status, and sexual status.

Community-based and school-based programs were defined as an intervention effort in which the primary locus of intervention was either in a community agency or a school. Of course, some programs are operated conjointly, and in most states, it is common for a school health service or a school-linked clinic to be viewed as a conjoint program between the community and the school (Schlitt, Ricket, Montgomery, & Lear, 1994). The question was how to code these programs. We determined the primary agent of the program and the delivery site. The delivery site was given weight in terms of coding. A common example found in coding the data was a hospital-sponsored program or a university-sponsored program offered in cooperation with a school system (i.e., Howard & McCabe, 1990). We called programs like this school-based because the programs were delivered in a school.

Interventions were categorized further into clinical and nonclinical programs to distinguish between prevention approaches that were delivered as a part of a medically oriented, school-based, or community-based health service and those originating in nonmedical programs, such as those offered through health education and social services. Skills and no skills were coded to indicate the presence of resistance, negotiation and decision-making training, and other life skills training programs that make use of social-cognitive and learning strategies to increase a particular behavior. Previous literature reviews indicated that these strategies may be effective (Kirby et al., 1994). For a program to be coded as including skills, the study had to indicate that program used this approach.

Both age and grade of study participants were collected to ascertain their maturation and developmental levels. Grade, however, was not included in the analysis because age was a better indicator of this variable. Sexual status was collected because previous studies found that the virgin or nonvirgin status of youths was an important intervening variable for the effects of sex education programs (Howard & McCabe, 1992; Kirby, Barth, Leland, & Fetro, 1991; Kirby, Waszak, & Ziegler, 1991).

Previous reviews of the outcome research distinguished studies along the lines of true experimental designs, quasi-experimental designs, and surveys (e.g., Kirby et al., 1994). We used the category of experimental-control group designs that also included quasi-experimental designs. Quasi-experimental designs were distinguished further by specific design, such as pretest, posttest, and posttest only designs. Critical reviews of meta-analyses indicate that the quality of experiments and research designs may contaminate the conclusions drawn from a meta-analysis (Kulik & Kulik, 1992). For this reason, we recommend that better-quality studies be analyzed both in conjunction with, as well as separate

from, the lower-quality studies to discover how the artifacts of poor study design affect the results (Krathwohl, 1993).

Beyond coding the studies for type of design, we developed a quality rating for each study that we reviewed. Studies were rated from 1 to 5, with 5 being a very high quality study and 1 a very low quality study. We used the following criteria, proposed by Gibbs (1989), to develop our rating form: type of research design, quality of the measurement system, and statistical procedures employed in a particular study. Two raters coded the studies for quality, and it was decided that the best method for resolving discrepancies between the two raters was to average the scores to obtain a quality rating. The deviation between the two raters was usually very small.

Because of the uneven distribution of the quality ratings, the scores were combined into a low, medium, and high ranking for the statistical analysis. The following rules were used: (a) average ratings of 2.5 or less were ranked low, (b) average scores between 3 and 4 were ranked medium, and (c) studies rated 4.5 or higher were ranked high.

Procedures

We followed the logic and the type of statistical methods recommended by Hedges (1982; Hedges & Olkin, 1985) and Rosenthal (1991). The three outcome variables—sexual activity, contraceptive use, and pregnancy rates—were analyzed in three separate meta-analyses, and effect sizes for each of the variables were kept independent. Studies were also examined for the effects of moderator variables, which included design, type of intervention, and type of program, as well as subject characteristics. In addition, the effect of the study design on outcomes was examined. When a study provided data on more than one behavioral outcome (i.e., sexual activity and contraceptive use), the results for each outcome were used in separate meta-analyses pertaining to that dependent variable (Rosenthal, 1991). When a study included measurements taken at two or more data points for one outcome variable (e.g., Eisen & Zellman, 1992; Howard & McCabe, 1992; Zabin et al., 1986a; Zabin, Hirsch, Smith, Streett, & Hardy, 1986b), the data were averaged. When different measurements were taken for a single outcome variable, they were combined up to the study level. For example, Zabin et al. (1986a, 1986b) included data for two questions about contraceptive use: the percentage of those who had ever used contraception, and the percentage of those who had used contraception at last intercourse. In this instance, effect sizes were computed for both measures and then averaged for a study-level effect size, which then was entered in the meta-analysis. This was done in order to maintain independent samples in the meta-analysis and to treat each study as the unit of analysis (Rosenthal, 1991). Occasionally, studies included examinations of more than one sample or intervention (i.e., Nicholson & Postrado, 1992). Each of these was included as a separate study for the meta-analysis.

Surveys made up a significant proportion of the studies on program effectiveness. In order to retain these posttest-only studies for the meta-analysis, we needed a comparison group so that an effect size could be calculated. For this purpose, we had national estimates matched for gender, age, and ethnicity as comparisons to calculate effect sizes. We employed data from national surveys such as Alan Guttmacher Institute age data from the National Survey of Family Growth Cycle IV.

Because most outcome variables were dichotomous and noncontinuous and often required a yes or no response from subjects, the data were presented in or transformed into proportions from frequencies. The simple differences between proportions, however, do not allow for the detection of differences in magnitude, given the location on the scale of proportions. Effect sizes were calculated for each dependent variable by first consulting the Cohen's (1977) Arcsine Transformation for Proportions Table given in Lipsey (1990) to yield arcsine transformations for treatment and control/comparison proportions. Control and comparison effect sizes were subtracted from the experimental effect size to produce an overall effect. Then the effect sizes were computed into correlations through a meta-analysis computer program known as XTOR (Veldman, 1995). Following this, the rs and ns for each study were entered into another computer program, RMETA, to produce a calculation of effect size. Most researchers recommend a transformation to Fisher's z before combining effect sizes in order to make the mean and variance independent (Rosenthal, 1991) and to attain a scale of equal units of detectability (Cohen, 1988). Therefore, we calculated the Fisher z transformation for each individual correlation r. Hedges and Olkin (1985) and Rosenthal (1991) believe that each of these z values should be weighted as a function of sample size because studies based on larger samples will

yield more precise estimates of effect size. We did this, then used the inverse function to convert the weighted z back to the metric of correlations. Finally, a fail-safe n was calculated for each of the three meta-analyses (Rosenthal, 1991). The fail-safe n tells us how many studies with opposite results would be necessary to nullify the result obtained through the meta-analysis.

To check our calculations, this procedure for obtaining measures of effect size and correlation coefficients from proportional data was recomputed using a second methodology for one of the measures (sexual activity). The D-Stat program (Johnson, 1989) uses a procedure in which differences between proportions are estimated by treating each as the mean of a distribution of 0s and 1s. The ensuing values matched those found in the earlier procedure.

An additional computer program, ANOVAN, was used for studying the effects of moderator variables. A chi-square test also was computed to conduct tests of homogeneity (Rosenthal, 1991). Homogeneity tests examine the possibility that systematic influences on the effect size produce significant variation in the measure of effect size. A significant test indicates the presence of significant variation, which may be explained systematically by specific characteristics of the studies, such as program characteristics (e.g., type of intervention, focus of intervention), subject characteristics (e.g., gender, age), or the design of the study.

Results

Effect sizes, z scores, type of design, and quality ratings are reported in the Appendix for the three outcome measures—sexual activity, contraceptive use, and pregnancy rates. We calculated effect size so that a positive effect size indicated changes in a positive direction. For example, if 20% of the control group engaged in sexual behavior and 10% of the treatment group did, the effect size would be positive for sexual behavior (i.e., the group undergoing treatment exhibited less sexual behavior, a positive outcome). However, if 20% of the control group used contraceptives and only 10% of the treatment group did, the effect size would be negative (i.e., the group undergoing treatment was actually using fewer contraceptives than the control group, a negative outcome).

As a rough guide to interpreting effect sizes, Cohen (1988) offers the labels "small, medium, and large." Other authors, however, caution against using such labels. They suggest that effect sizes be evaluated considering the specific goals of a program or treatment (Glass, McGaw, & Smith, 1981). For the three meta-analyses, the significance of the weighted z value tells us if there is a significant difference between the treatment and control groups, the pre- and posttest scores, or the program and national estimates.

Sexual Activity

Altogether, 17 studies reported measures of sexual activity. Twenty-four independent outcomes were reported because some studies included more than one intervention or sample group. A combined effect size (*ES*) for those studies, $ES = .0110$, is nonsignificant. Effect sizes range from –0.424 to 0.571. Only five studies showed a low to moderate positive effect size, meaning a decrease in sexual activity, for the outcome variable. Study results also yielded a homogeneity chi-square of 80.24, indicating the study results varied widely and warranting follow-up analyses with the moderator variables. The fail-safe n for the nonsignificant results of the sexual activity measures was 63. This indicates that 63 studies with opposite positive effects would need to be added to the 24 outcomes before the cumulative results became significant.

The results of the analyses of moderator variables indicated that the following variables had no power to explain the variability in the overall meta-analysis: locus of intervention (community vs. school-based), type of program (clinic vs. nonclinic), type of intervention (abstinence training vs. sexual education vs. contraceptive use), focus of intervention (skills vs. no skills), design of study (experimental-control vs. pretest or posttest), quality of study (high vs. medium vs. low), and gender (males vs. females). Sexual status (virgins vs. nonvirgins at the start of study) and socioeconomic status were not reported consistently enough in the studies to draw conclusions from these data.

However, age and ethnicity did show significant differences between the groups, with the findings yielding mixed results. Studies that examined the entire span of teenage years (11–20) had a significant positive effect size ($ES = .13$), indicating a decrease in sexual activity. Studies that looked only at 11–15 year olds or 16–20 year olds did not show differences in sexual activity. When data were available for different age groups within studies, we entered these data points separately into a second analysis, which indicated there were

no differences in program effectiveness for sexual activity for younger versus older teens.

The same pattern of results emerged for ethnicity. There was no difference in program effectiveness for studies that had mostly White or mostly Black samples (at least 70%). For the studies with Latino samples, the effect size was negative and nonsignificant (*ES* = –.1153). Again, we completed a second analysis and separated studies that gave information about more than one group. The difference between ethnic groups became nonsignificant.

Contraceptive Use

Sixteen studies, totaling 22 different outcomes, reported measures of contraceptive use. A combined effect size for those studies yielded a highly significant, low effect size, *ES* = .27. Effect sizes ranged from –0.2945 to 1.151. The chi-square test for homogeneity was 470.68, indicating considerable variation and warranting a follow-up analysis with moderator variables. The fail-safe *n* for the results of the measures of contraceptive use was 5,624. This indicates that 5,624 studies with null-summing effects would need to be added to the 22 outcomes before the cumulative results became nonsignificant.

Analyses of the moderator variable indicated that community-based programs resulted in increased contraceptive use (*ES* = .6062, medium effect size) over school-based programs (*ES* = .1195), although both resulted in significant positive effects. Contraceptive use also improved more in clinics (*ES* = .3355, low effect size) than in nonclinical programs (*ES* = .0711). Programs that emphasized contraceptive distribution and knowledge building (*ES* = .3313, low effect size) were more effective than programs that emphasized solely sex education (*ES* = .0638). The no-skills treatment (*ES* = .3301), compared with the skills approach (*ES* = .0557), improved contraceptive use in all types of programs. The no-skills approach seemed to be more effective for community-based programs (*ES* = .6196) than for school-based programs (*ES* = .1461). A follow-up analysis revealed, however, that this finding was mediated by the fact that clinic programs in schools and communities were more frequently coded as not providing skills training than nonclinic types of programs (*r* = .886). Therefore, this result is believed to reflect the effect of clinic programs versus nonclinic programs, versus than the effect of an emphasis on skills training over no skills training.

Because one of the determinants of our confidence in the evaluation of program effectiveness is the design of the study, both the type of design and the quality rating were evaluated in relationship to the overall effects of the program and contraceptive use outcomes. Findings indicate that the effects remained positive for experimental-control designs (*ES* = .1126) and pretest and posttest designs (*ES* = .5603), but yielded a small negative effect (*ES* = –.0002) for comparisons of posttest data and national data. Experimental-control designs yielded significantly lower effect sizes than noncontrolled pretest and posttest designs, yet both were significantly different from zero. When we examined quality ratings, we found that all studies, regardless of rating, yielded a positive effect size. The quality rating did, however, make a large difference in the size of the effect obtained. Studies rated high quality obtained significance but with the lowest effect size (*ES* = .0762), compared with studies with medium ratings, which obtained medium effect sizes (*ES* = .4719) and studies with low ratings, which obtained a low, yet significant, effect size (*ES* = .1907).

Subject characteristics also emerged as salient in relationship to the contraceptive use outcome. Older adolescents (aged 15–20; *ES* = .2210) appear to perform better on measures of contraceptive use than younger adolescents (aged 11–14; *ES* = .1164). Gender differences indicated that studies with only females (*ES* = .1497) and studies with mixed male and female groups (*ES* = .2878) demonstrate small, positive effect sizes on the contraceptive use variable. The difference between the two groups is significant and favors the mixed-gender group. We entered studies that provided data for both male and female groups into a second analysis. This analysis indicated no significant difference between males and females (*ES* = .1078 for males; *ES* = .1731 for females).

Ethnic comparisons indicate that Latinos (*ES* = .7870) make up the ethnic group most likely to perform best on the measures of contraceptive use following intervention. African Americans yielded only a small, positive effect size (*ES* = .1667), and a mixed category of African Americans and European Americans obtained a small, negative effect size (*ES* = –.1390), compared with the Latino group. The samples yielded no groups made up of only European Americans or only Native Americans. Socioeconomic status was not reported with enough consistency in the studies to draw

conclusions from this data. As might be expected, exclusively nonvirgin groups appeared for the contraceptive use outcome variable. A mixed group of nonvirgins and virgins (*ES* = .4140) and a group of all nonvirgins (*ES* = .1456) indicate that the mixed group yielded a medium, positive effect size, compared with the small, positive effect size obtained by the nonvirgin group.

Pregnancy Rates

Fifteen studies consisting of 25 independent outcomes reported measures of pregnancy rates. A combined effect size for those studies yielded a small but significant effect size, *ES* = .153. Effect sizes ranged from –.2517 to .4735. Study results yielded a homogeneity chi-square of 238.38, indicating that study results varied immensely. The fail-safe *n* for the nonsignificant results of the measures of sexual activity was 3,339—which means 3,339 studies with null-summing effects would need to be added to the 25 outcomes before the cumulative results became nonsignificant. Although less significant than contraceptive use, program characteristics did impact pregnancy rates. Similar to the findings for contraceptive use, we found that community-based programs (*ES* = .2753) were more effective than school-based programs (*ES* = .0920). Clinic programs (*ES* = .2030) also have larger effect sizes than nonclinic programs (*ES* = .0718). Programs emphasizing contraceptive use and distribution (*ES* = .1996) performed better than abstinence-based programs (*ES* = .0623) and sex education programs without contraceptive knowledge building (*ES* = .0818). In addition, no-skills programs (*ES* = .1798) yielded larger effect sizes than skills programs (*ES* = .0206). However, as reported above, the clinic-based programs appear responsible for this finding (r = .603).

The quality rating of the study and the design of the study, two related variables, had an effect on pregnancy rates similar to the one for measures of contraceptive outcomes. The type of design considerably affected the results. The study design of posttest only (national surveys) reported the highest effect size for pregnancy rates (*ES* = .2700), compared with pretest and posttest studies (*ES* = .0489) and experimental-control studies (*ES* = .0998). Low (*ES* = .2035) and medium quality studies (*ES* = .2474) had positive effect sizes, whereas high quality studies showed a slight negative effect (*ES* = –.0662).

Regarding subject-related characteristics coded as moderator variables, groups of females performed better on measures of pregnancy outcomes (reduced pregnancies, (*ES* = .3112) than groups of mixed females and males (*ES* = .0761). A second analysis was completed when studies reported data for both male and female groups. These results indicated that females (*ES* = .2431) had significantly lower pregnancy rates after participating in a program, but male behavior (i.e., getting someone pregnant, *ES* = –.0172) did not change significantly. Age affected effect size. The older group (15 years old and older) obtained a slightly positive effect size (*ES* = .1255), and the younger group (younger than 15 years) obtained a slightly negative effect (*ES* = –.0727). No differences for ethnicity were demonstrated between African Americans and European Americans, although the number of African American participants was much higher in the studies.

Socioeconomic and sexual status were not reported in many studies. Therefore, the following results are based on very low numbers. In general, program participants tended to have low socioeconomic status. Programs that were primarily oriented to low groups with socioeconomic status performed significantly better than programs with mixed socioeconomic status groups (*ES* = .1942 for low; *ES* = .0527 for mixed). The variable of sexual status showed one significant difference. If a program had all nonvirgins at the start, participants were more successful at not getting pregnant (*ES* = .1215) than if the participants were both virgins and nonvirgins (*ES* = –.0557).

Discussion

Results of the meta-analysis indicate that pregnancy prevention programs fail to affect the sexual activity of adolescents, a finding often reflected in the narrative reviews of the research literature and echoed by the critics of sex education (Whitehead, 1994). This finding differs, however, from the findings of Frost and Forrest (1995), who concluded after reviewing the nonstandardized effect sizes of five school-based programs that these programs were delaying the sexual activity of participants. Maybe specific programs such as the ones evaluated in the Frost and Forrest review are able to affect the sexual behavior of some adolescents. Looking at these five studies individually, as Kirby et al. (1994) did, would lead to this conclusion. But without further replications with larger samples, the claim could not be made that

programs were affecting adolescents' sexual behavior. The positive findings concerning sexual activity do not hold up when these five studies are synthesized using more rigorous meta-analytical methods, along with the inclusion of a larger pool of diverse studies from both school and community settings. The fail-safe *n* indicates that 63 studies with positive effects are needed to reverse the present findings.

Prevention programs are more successful in affecting contraceptive use and pregnancy rates. These findings constitute more optimistic outcomes than previous narrative reviews, which reported inconsistent, inconclusive, or negative findings concerning contraceptive use and pregnancy rates (Dawson, 1986; Hofferth, 1991; Kirby, 1992; Kirby et al., 1994). Findings concerning contraceptive also support the Frost and Forrest (1995) meta-analysis, which concluded that pregnancy prevention programs are able to increase the contraceptive use of sexually active teens. In contrast, findings concerning pregnancy rates contradict the results of that study, results that indicate that programs have little impact on pregnancy rates. We hypothesize that community-based clinic programs may be mostly responsible for the differences in findings between our study and the Frost and Forrest study. We also believe that with the current political climate and the increased public concern about teens giving birth, our findings of positive effects of programs on rates pregnancy are critical in demonstrating the utility of programs.

The results of our meta-analysis pinpointed characteristics of programs and subject variables of considerable importance. Kirby et al. (1994) emphasized the importance of determining the effectiveness of prevention programs in relationship to their locus of intervention (school vs. community). Findings from our study indicate that contraceptive use improves in community-based programs more often than in school-based programs. Community-based programs achieve a positive, medium effect size. This finding was replicated on a smaller scale for the pregnancy rate outcomes.

Our meta-analysis further examined clinical programs and nonclinical alternatives to determine which is more effective. Previous qualitative reviews of the literature indicate that clinical programs are more effective than nonclinical programs in increasing the contraceptive use of adolescents (Dawson, 1986). Hofferth (1991) also reported that community-based clinics, not school-based clinics, reduce pregnancy rates. These claims were supported by our meta-analysis. Clinic-based services more effectively influenced contraceptive behavior outcomes than nonclinical programs. Clinic programs had more impact on pregnancy rates, although the overall effectiveness of programs on pregnancy rates disappears when the research design is controlled for quality. The effects on the contraceptive use outcome are also diminished after controlling for quality of design, although they still remain positive.

In addition, for pregnancy outcome measures, contraceptive knowledge-building programs and contraceptive distribution programs are more effective than other sex education programs, including ones that stress abstinence. Contraceptive knowledge-building programs and contraception distribution are often controversial and not accepted by many parents, school districts, and communities. A common oppositional ideology to the programs is that educating adolescents about contraceptives and making contraceptives available will increase the likelihood that teenagers will participate in sexual intercourse and get pregnant. However, as Kirby et al. (1994) state in their narrative review, studies indicate that such programs do not encourage sexual involvement, even among participants 14 years old and younger. This seems to be supported by the meta-analysis, which found that contraceptive programs were the most effective programs, positively affecting two of the outcome measures.

Another concern involves the mixed message that is given when schools teach abstinence from sexual intercourse in juxtaposition with contraceptive advice. Many school personnel, school board members, community leaders, and other concerned people fear that this may lead to confusion among adolescents and may reduce the credibility of the program. The concern about mixed messages has been addressed by other evaluators, who have argued that a program can both reduce rates of sexual activity and improve rates of contraceptive use (Frost & Forrest, 1995; Howard & McCabe, 1990). When we examined only educational programs, then all programs, we found mixed results as to whether both sexual activity and contraceptive use were affected positively. In a comparative analysis of the sexual behavior of teens from several developed countries, Jones, Namerow, and Philliber (1982) argue that the "two-pronged" approach to pregnancy prevention in this country, involving both abstinence and contraceptive use, divides the effectiveness of efforts. The authors present evidence that other de-

veloped countries' programs are more effective in influencing teen sexual behavior by promoting the use of contraceptives.

Our study shows that school-based clinics were not as effective as community-based clinics, although school-based clinics did affect contraceptive use more effectively than other school-related sex education programs. One explanation for the better performance of the community-based clinics is that school-based clinics have a low percentage of contraceptive distribution, compared with community-based, family planning clinics. One report indicates that fewer than 20% of school-based clinics are reported to actually distribute contraceptives, and only 28% of these write prescriptions for birth control pills (Kirby et al., 1994). Thus, although contraceptive distribution has proven effective, most school-based clinics are prohibited from this practice.

In this study, we were not able to clarify the effectiveness of skills training interventions, compared with interventions that did not teach skills, because the coding of this variable seemed to be confounded by the fact that clinic programs, responsible for the majority of positive changes in outcome variables, were usually coded as offering no skills training. It is unfortunate that we were not able to investigate this training variable because approaches to sex education that are social-cognitive and learning-based are more effective than other approaches to sex education (Kirby et al., 1994). Other studies will have to clarify the specific learning strategies that community-based clinics may be using in their work that yields positive results on adolescent behavior.

Females performed better on measures of pregnancy outcome (reduced pregnancies and childbirths) than males. At the same time, mixed groups of females and males who participated in programs increased their use of contraceptives more than groups of females only. However, there is a shortage of young men participating in programs. It seems important to include more males and also to find new ways to teach them responsible sexual behavior.

Age proved to be an interesting variable. Results indicate that older adolescents (aged 15–19) perform better on measures of contraceptive use than those 14 years old and younger. Younger women are also more likely to get pregnant. These results are consistent with previous research that found that older female adolescents are more likely to use contraceptives effectively than younger adolescents. The Alan Guttmacher Institute (1994a) reported that younger females rely on less effective methods of contraception, such as condoms and withdrawal, and that young adolescents are likely to use contraception inconsistently.

Even though contraceptive programs seem to work best with older, sexually active teens, this does not mean that younger, sexually active teens cannot be taught effective use of contraceptives (see Howard & McCabe, 1990; Kirby et al., 1991, 1994). The critical point may be how to work with younger teens in the use of contraceptives. For example, it may be important to work with males' discomfort with condoms and to attack myths and erroneous beliefs—such as oral contraceptives are harmful or a person cannot get pregnant during first sexual intercourse—that are common to this group. The inability to anticipate sex ahead of time or to set clear goals about sexual involvement may further inhibit effective contraceptive use. Previous narrative reviews recommend programs designed along the dimensions of adolescent development, age, and sexual experience of the participants (Kirby et al., 1994). Comprehensive, age-phased, developmental approaches such as the one suggested by Nicholson and Postrado (1992) may be appropriate. Age differences found in our meta-analysis appear to support this conclusion, and further research is needed to determine the effectiveness of varying approaches with younger adolescents.

Program developers also may want to take into consideration that most young teens are not sexually active. For example, 84% of 13 year olds, 77% of 14 year olds, and 70% of 15 year olds have not had sexual intercourse (Alan Guttmacher Institute, 1994b). Because the majority of young teens is not sexually active, career-focused programs that encourage setting goals and postponing sexual involvement in favor of the pursuit of higher aspirations may be important strategies for working with them. Although the meta-analysis did not specifically focus on the effectiveness of career programs in relationship to contraceptive use and pregnancy prevention, given the developmental differences between younger and older adolescents, this may be a useful area of inquiry. Career programs, of course, should always be offered with education about contraceptives, so that teens will be prepared for sexual activity.

Other moderator analyses showed no consistent differences between African Americans and European Americans. Latinos were shown to perform best on contraceptive use measures following intervention. Overall, the number of African

American participants was much higher throughout the studies than participants from other ethnic groups. Also, most participants came from lower socioeconomic groups, which precluded meaningful comparisons among ethnic and socioeconomic groups. Most programs included high concentrations of teenagers from ethnic and low socioeconomic groups, and this may indicate how social policies are formed regarding funding issues. These groups may be perceived as being at higher risk for adolescent pregnancy than are European Americans. A perceived bias may be in effect here against early childbearing within ethnic minority groups, and programs may be enacted to prevent its occurrence. A question for further study is what kinds of results may be achieved with European American groups.

The meta-analysis addressed another important area often neglected by narrative reviews. It involves the quality ratings of studies and their effect on outcomes. Results of the meta-analysis indicate that higher quality studies do not yield the same positive results on contraceptive use and pregnancy rate outcomes. The quality of research in the field of pregnancy prevention has been extremely poor (Dawson, 1986; Frost & Forrest, 1995; Hofferth, 1991; Miller, Card, Paikoff, & Peterson, 1992). We agree that the state of the research is not optimal. It was difficult to find studies appropriate for a meta-analysis because the overall state of the research did not provide the type or the quality of data necessary. However, this is not uncommon. For example, research on the effectiveness of alternative schools and dropout programs is in a similar state (Franklin, 1992).

Limitations of this study include those usually associated with meta-analyses. Three sets of bias may preclude definite conclusions. First, some publications may have been overlooked in the collection of studies. Second, some findings may have been excluded because the studies were rejected for publication by journal review boards. Third, viable studies may have been left out because researchers neglected to get them published. This so called "File Drawer" problem (Rosenthal, 1984) is difficult to solve. Researchers can never be sure that they have sampled every study. We are certain we did not find and include every possible study. We did, however, calculate a fail-safe *n*, a rough estimate to help researchers gauge how many studies with opposite effects would be necessary to reverse the findings of the meta-analysis. A large number of studies would be necessary to change the results that we have reported, which helps increase our confidence in the findings.

Another issue for the meta-analysis is that the types of studies included may be determined by the statistics presented in the studies. Unfortunately, it was often difficult to include studies with advanced statistics, such as log-linear and multiple regression analysis, because the numbers reported in the studies were not congruent with the types of elementary statistics needed to calculate effect sizes. Every effort was made to include such studies if the researchers could use secondary analysis of the data to produce the statistics needed.

A specific issue for this study involved the partial use of data within studies or changing the study designs in order to use the data. For example, some pretest and posttest, experimental, and quasi-experimental designs were translated into posttest-only designs in order to use them. These design transformations were necessary when data for control or comparison groups were not clearly presented. It would have been desirable to have contacted researchers to collect these data, but the constraints of time and resources did not make this possible. Despite these limitations, our study represents a needed quantitative analysis of the literature on studies of pregnancy prevention outcomes. We hope that other researchers will conduct additional meta-analytical reviews and build on our work.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Alan Guttmacher Institute (1994a). *Sex and America's teenagers.* New York: Author.

Alan Guttmacher Institute (1994b). *Teenage reproductive health in the United States.* New York: Author.

Bangert-Drowns, R. L. (1986). Review of developments in meta-analytic method. *Psychological Bulletin, 99,* 388-399.

Beck, J. G., & Davies, D. K. (1987). Teen contraception: A review of perspectives on compliance. *Archives of Sexual Behavior, 16,* 337-376.

*Berger, D. K., Perez, G., Kyman, W., Perez, L., Garson, J., Menendez, M., Bistritz, J., Blanchard, H., & Dombrowski, C. (1987). Influence of family planning counseling in an adolescent clinic on sexual activity and contraceptive use. *Journal of Adolescent Health Care, 8,* 436-440.

*Christopher, F. S., & Roosa, M. W. (1990). An evaluation of an adolescent pregnancy prevention program: Is "Just say no" enough? *Family Relations, 39,* 68-72.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York: Academic Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cook, T. D., & Leviton, L. C. (1980). Reviewing the literature: A comparison of traditional methods with meta-analysis. *Journal of Personality, 48,* 449-472.

Dawson, D. A. (1986). The effects of sex education adolescent behavior. *Family Planning Perspectives, 18,* 162-170.

*Edwards, L., Steinman, M., Arnold, K., & Hakanson, E. (1980). Adolescent pregnancy prevention services in high school clinics. *Family Planning Perspectives, 12,* 6-14.

*Eisen, M., & Zellman, G.L. (1987). Changes in the incidence of sexual intercourse of unmarried teenagers following a community-based sex education program. *Journal of Sex Research, 23,* 527-533.

*Eisen, M., & Zellman, G. L. (1992). A health beliefs field experiment: Teen talk. In B. C. Miller, J. J. Card, R. L. Paikoff, & J. L. Peterson (Eds.), *Preventing adolescent pregnancy* (pp. 220-264). Newbury Park, CA: Sage.

*Eisen, M., Zellman, G. L., & McAlister, A. (1985). A health belief model approach to adolescents' fertility control: Some pilot program findings. *Health Education Quarterly, 12,* 185-210.

Franklin, C. (1992). Alternative school programs for at-risk youth. *Social Work in Education, 14,* 239-251.

*Freeman, E. W., Rickels, K., Mudd, E. B. H., & Huggins, G. R. (1982). Never-pregnant adolescents and family planning programs: Contraception, continuation, and pregnancy. *American Journal of Public Health, 72,* 815-822.

Frost, J. J., & Forrest, J. D. (1995). Understanding the impact of effective teenage pregnancy programs. *Family Planning Perspectives, 27,* 188-195.

*Furstenberg, F. F., Jr., Herceg-Baron, R., Shea, J., & Webb, D. (1984). Family communication and teenagers' contraceptive use. *Family Planning Perspectives, 16,* 163-170.

*Furstenberg, F. F., Jr., Moore, K. A., & Peterson, J. L. (1985). Sex education and sexual experience among adolescents. *American Journal of Public Health, 75,* 1331-1332.

Gibbs, L. (1989). Quality of Rating Form: An instrument for synthesizing evaluation studies. *Journal of Social Work Education, 25,* 55-67.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5,* 3-8.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.

Hedges, L. V. (1982). Estimation of effect sizes from a series of independent experiments. *Psychological Bulletin, 92,* 490-499.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

*Herceg-Baron, R., Furstenberg, F. F., Shea, J., & Harris, K. M. (1986). Supporting teenagers' use of contraceptives: A comparison of clinic services. *Family Planning Perspectives, 18,* 61-66.

Hofferth, S. L. (1991). Programs for high-risk adolescents. *Evaluation and Program Planning, 14,* 3-16.

*Howard, M., & McCabe, J. A. (1990). Helping teenagers postpone sexual involvement. *Family Planning Perspectives, 22,* 21-26.

*Howard, M., & McCabe, J. A. (1992). An information and skills approach for younger teens: Postponing sexual involvement program. In B. C. Miller, J. J. Card, R. L. Paikoff, & J. L. Peterson (Eds.), *Preventing adolescent pregnancy* (pp. 83-109). Newbury Park, CA: Sage.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of Meta-analysis.* Newbury Park, CA: Sage.

Iverson, B. K., & Levy, S. R. (1982). Using meta-analysis in health education research. *The Journal of School Health, 52,* 234-239.

Johnson, B. T. (1989). *D-Stat: Software for the meta-analytic review of research literatures.* Hillsdale, NJ: Erlbaum.

*Jorgensen, S. R., Potts, V., & Camp, B. (1993). Project taking charge: Six-month follow-up of a pregnancy prevention program for early adolescents. *Family Relations, 42,* 401-406.

Kirby, D. (1992). School-based programs to reduce sexual risk-taking behaviors. *Journal of School Health, 62,* 280-287.

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

*Kirby, D., Harvey, P. D., Claussenius, J., & Novar, M. (1989). A direct mailing to teenage males about condom use: Its impact on knowledge, attitudes, and sexual behavior. *Family Planning Perspectives, 21,* 12-18.

Kirby, D., Short, L., Collins, J., Rugg, D., Kolbe, L., Howard, M., Miller, B., Sonenstein, F., & Zabin, L. (1994). School-based programs to reduce sexual risk behaviors: A review of effectiveness. *Public Health Reports, 109,* 339-360.

*Kirby, D., Waszak, C., & Ziegler, J. (1991). Six school-based clinics: Their reproductive health services and impact on sexual behavior. *Family Planning Perspectives, 23,* 6-16.

Krathwohl, D. R. (1993). *Methods of educational and social science research: An integrated approach.* New York: Longman.

Kulik, J. A., & Kulik, C. C. (1992). Meta-analysis: Historical origins and contemporary practice. In B. Thompson (Ed.), *Advances in social science methodology, Vol. 2,* (pp. 53-79). Greenwich, CT: JAI Press.

Light, R. J., & Smith, P. V. (1971). Accumulating evidence: Procedures for resolving contradictions among different research studies. *Harvard Educational Review, 41,* 429-471.

Lipsey, M. W. (1990). *Design sensitivity: Statistical power for experimental research.* Newbury Park, CA: Sage.

*Marsiglio, W., & Mott, F. (1986). The impact of sex education on sexual activity, contraceptive use, and premarital pregnancy among American teenagers. *Family Planning Perspectives, 18,* 151-162.

Miller, B. C., Card, J. J., Paikoff, R. C., & Peterson, J. C. (1992). *Preventing adolescent pregnancy: Model programs and evaluation.* Newbury Park, CA: Sage.

*Mudd, E. H., Dickents, H. O., Garcia, C. R., Rickels, K., Freeman, E., Huggins, G. R., & Logan, J. J. (1978). Adolescent health services. *American Journal of Orthopsychiatry, 48,* 495-504.

*Namerow, P. B., & Jones, J. E. (1982). Ethnic variation in adolescent use of a contraceptive service. *Journal of Adolescent Health Care, 3,* 165-172.

*Namerow, P. B., Philliber, S. G., & Hughes, M. (1983). Follow-up of adolescent family planning clinic users. *Family Planning Perspectives, 15,* 172-176.

*Nicholson, H. J., & Postrado, L. T. (1992). A comprehensive age-phased approach: Girls Incorporated. In B. C. Miller, J. J. Card, R. L. Paikoff, & J. L. Peterson (Eds.), *Preventing adolescent pregnancy* (pp. 110-133). Newbury Park, CA: Sage.

*Philliber, S. & Allen, J. P. (1992). Life options and community service: Teen outreach program. In B. C. Miller, J. J. Card, R. L. Paikoff, & J. L. Peterson (Eds.), *Preventing adolescent pregnancy* (pp. 139-155). Newbury Park, CA: Sage.

*Postrado, L. T., & Nicholson, H. J. (1992). Effectiveness in delaying the initiation of sexual intercourse of girls aged 12-14: Two components of the Girls Incorporated Preventing Adolescent Pregnancy Program. *Youth and Society, 23,* 356-378.

*Roosa, M. W., & Christopher, F. S. (1990). Evaluation of an abstinence-only adolescent pregnancy prevention program: A replication. *Family Relations, 39,* 363-367.

Rosenthal, R. (1984). *Meta-analytic procedures for social research.* Newbury Park, CA: Sage.

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (Rev. ed.). Newbury Park, CA: Sage.

*Schinke, S. P., Blythe, B., & Gilchrist, L. (1981). Cognitive behavioral prevention of adolescent pregnancy. *Journal of Counseling Psychology, 28,* 451-454.

*Schinke, S. P., Blythe, B. J., Gilchrist, L. D., & Burt, G. A. (1981). Primary prevention of adolescence. *Social Work with Groups, 4,* 121-135.

Schlitt, J. J., Ricket, K. D. Montgomery, L. L. & Lear, J. G. (1994, October). *State initiatives to support school-based health centers: A national survey.* Paper presented at the annual meeting of the National Association of Social Workers, Memphis, TN.

*Shea, J. A., Herceg-Baron, R., & Furstenberg, F. F. (1984). Factors associated with adolescent use of family planning clinics. *American Journal of Public Health, 74,* 1227-1230.

*Thomas, B. H., Mitchell, A., Devlin, M. C., Goldsmith, C. H., Singer, J., & Watters, D. (1992). Small group sex education at school: The McMaster Teen Program. In B. C. Miller, J. J. Card, R. L. Paikoff, & J. L. Peterson (Eds.), *Preventing adolescent pregnancy* (pp. 28-52). Newbury Park, CA: Sage.

Veldman, D. J. (1995). XTOR meta-analysis program. Department of Psychology, University of Texas at Austin.

*Vincent, M. L., Clearie, A. F., & Schluchter, M. D. (1987). Reducing adolescent pregnancy through school and community-based education. *Journal of the American Medical Association, 257,* 3382-3386.

Whitehead, B. D. (1994, October). The failure of sex education. *The Atlantic Monthly, 274,* 55-80.

Whitley, B. E., & Schofield, J. W. (1986). A meta-analysis of research on adolescent contraceptive use. *Population and Environment, 8,* 173-204.

Zabin, L. S. (1992). School-linked reproductive health services: The Johns Hopkins program. In B. C. Miller, J. J. Card, R. L. Paikoff, & J. L. Peterson (Eds.), *Preventing adolescent pregnancy: Model programs and evaluations* (pp. 156-184). Newbury Park, CA: Sage.

Zabin, L. S., Hardy, J. B., Streett, R., & King, T. M. (1984). A school-, hospital- and university-based adolescent pregnancy prevention program: A cooperative design for service and research. *The Journal of Reproductive Medicine, 29,* 421-426.

*Zabin, L. S., Hirsch, M. B., Smith, E. A., Streett, R. & Hardy, J. B. (1986a). Adolescent pregnancy-prevention program: A model for research and evaluation. *Journal of Adolescent Health Care, 7,* 77-87.

*Zabin, L. S., Hirsch, M. B., Smith, E. A., Streett, R., Hardy, J. B. (1986b). Evaluation of a pregnancy prevention program for urban teenagers. *Family Planning Perspectives, 18,* 119-126.

Zelnik, M., & Kim, Y. J. (1982). Sex education and its association with teenage sexual activity, pregnancy and contraceptive use. *Family Planning Perspectives, 14,* 117-125.

APPENDIX

Author	Locus	Design and Quality Ratings*	Outcomes Used in Meta-Analysis†		
			n	*ES*	*z*
Berger et al. (1987)	Community-based (municipal outpatient adolescent health) clinic; low socioeconomic area of New York	Pretest, posttest/2.5	Sexual activity 383 Contraceptive use 142	0.04 1.03	0.02 0.49
Christopher & Roosa (1990) Christopher & Roosa (1990)	School-based (Adolescent Family Life Act); low-income neighborhoods Tempe, AZ	Quasi-experimental; pretest, posttest/4	Sexual activity 320 Sexual activity 361	–0.31 0.00	0.15 0.00
Edwards et al. (1980)	School-based clinic	Posttest only, including analysis of clinic records and tracking fertility rates/2.5	Pregnancy rates 403	0.14	0.07
Eisen et al. (1985)	Community-based (university) clinic; Austin and San Marcos, TX	Pretest, posttest, follow-up/13.5	Sexual activity 120 Contraceptive use 120	0.01 0.03	0.01 0.02
Eisen & Zellman (1987)		Pretest, posttest/2	Sexual activity 120	–0.42	–0.21
Eisen & Zellman (1992) Posttests only used in meta-analysis	Mainly community-based program; TX, CA	Quasi-experimental; pretest, posttest, follow-up/3	Sexual activity 888	0.28	0.14
Freeman et al. (1982) Posttests used for pregnancy outcome	Community-based program (university hospital in Philadelphia)	Pretest, posttest/2.5	Sexual activity 321 Contraceptive use 314 Pregnancy rates 346	0.04 0.77 –0.01	0.02 0.38 –0.00
Furstenberg et al. (1985)	National survey of school-based programs	Quasi-experimental comparison of participants who had sex education with those who had not/3	Sexual activity 469	0.22	0.11
Herceg-Baron et al. (1986) Shea et al. (1984) Posttest only used in meta-analysis	Community-based program; family planning agencies, health clinics in Philadelphia	Experimental; participants randomly assigned to treatment or control (regular services) conditions; pretest, posttest, follow-up/2	Contraceptive use 328 Pregnancy rates 540 Contraceptive use 359	–0.10 0.03 –0.29	–0.47 0.01 –0.15
Furstenberg et al. (1984) Pretest and posttest used in meta-analysis	Community-based program; family planning agencies, health clinics in Philadelphia	Experimental; participants randomly assigned to treatment or control (regular services) conditions; pretest, posttest, follow-up/3	Contraceptive use 274	0.01	0.00
Howard & McCabe (1990) Howard & McCabe (1992)	School-based program developed by hospital; (Atlanta)	Quasi-experimental; matched school design; pretest, posttest/3 3.5	Sexual activity 364 Pregnancy rates 238 Sexual activity 510	0.42 0.07 0.07	0.19 0.19 0.03

APPENDIX (CONTINUED)

Author	Locus	Design & Quality Ratings*	Outcomes Used in Meta-Analysis†		
			n	*ES*	*z*
Jones et al. (1982)	Community-based (hospital) program; Washington Heights, NY	Posttest only/2	Pregnancy rates		
			3078	0.27	0.14
Namerow et al. (1983)		Posttest/1.5	Contraceptive use		
			178	0.66	0.30
Namerow & Jones (1982)		Pretest, posttest/3	Contraceptive use		
			3858	0.80	0.39
Posttest only for pregnancy outcome			Pregnancy rates		
			3858	0.31	0.15
Jorgensen et al. (1993)	School-based (university) program; DE and MS	Quasi-experimental; pretest, posttest, follow-up/4	Sexual activity		
			91	0.57	0.28
Kirby et al. (1991)	School-based program; rural and urban areas of California	Quasi-experimental; pretest, posttest, follow-up/2.5	Sexual activity		
			758	0.57	0.28
			Contraceptive use		
			758	–0.08	–0.04
		3.5	Pregnancy rates		
			758	–0.09	–0.04
Kirby et al. (1989)	Community-based program; low-income communities nationwide	Quasi-experimental; participants randomized to treatment and control conditions/4	Sexual activity		
			2017	–0.06	–0.03
Kirby et al. (1991) Multiple outcomes derived from a single study	School-based programs; in Gary, IN; San Francisco; Muskegon, MI; Jackson, MS; Quincy, FL; Dallas, TX	At four schools, quasi-experiment with matched schools as comparison groups/4.5	Sexual activity		
			824	–0.11	–0.05
			1360	0.1	–0.05
			430	–0.04	–0.02
		At two schools, pretest, posttest design used/4.5	587	0.03	0.01
			1295	0.04	0.02
			1274	0.03	0.01
			Contraceptive use		
			1035	–0.19	–0.88
			221	0.17	0.09
			904	0.03	0.01
			687	0.15	0.07
			1078	0.33	0.15
			574	0.12	0.58
			Pregnancy rates		
			1100	–0.04	–0.02
			691	–0.08	–0.16
			895	0.05	0.02
			212	0.00	0.00
			556	0.07	0.03
			958	–0.09	–0.04
Marsiglio & Mott (1986)	National survey	Quasi-experimental with teens who had completed programs and those who had not/3	Contraceptive use		
			477	0.19	0.09
Mudd et al. (1978)	School-based or school-linked clinic providing knowledge of contraceptive services	Posttest/2.5	Pregnancy rates		
			161	0.28	0.14

APPENDIX (CONTINUED)

Author	Locus	Design & Quality Ratings*	Outcomes Used in Meta-Analysis†		
			n	*ES*	*z*
Nicholson & Postrado (1992) Posttests only used in meta-analysis	Community-based program (Girls Incorporated) where adolescent pregnancy rates were higher than national average	Quasi-experimental; compared participants who volunteered for treatment with those who did not/4	Pregnancy rates		
			15	–0.25	–0.13
			25	–0.17	–0.09
			9	–0.03	–0.17
			84	0.11	0.06
			Sexual activity		
Postrado & Nicholson (1992)			412	0.24	0.09
			412	0.02	0.01
Philliber & Allen (1992)	School-based program (usually in collaboration with Junior League)	Quasi-experimental with treatment and comparison groups; pretest, posttest/3	Pregnancy rates		
			302	0.26	0.13
			986	0.11	0.55
Schinke, Blythe, & Gilchrist (1981)	School-based program (university); Seattle	Solomon 4-group design/4	Contraceptive use		
			36	1.15	0.55
Schinke, Blythe, Gilchrist, & Burt (1981)	School-based program (university); Seattle	Experimental. Participants randomly assigned to treatment and control	Contraceptive use		
			53	0.83	0.41
Thomas et al. (1992)	School-based program; Hamilton, Ontario, Canada	Experimental. Schools randomly assigned to treatment	Sexual activity		
			2570	–0.10	–0.05
			Pregnancy rates		
			2570	0.08	–0.03
			Contraceptive use		
			2570	–0.07	–0.03
Vincent et al. (1987)	Community-based program; rural South Carolina	Quasi-experimental with other counties as matched comparison groups/2.5	Pregnancy rates		
			740	0.47	0.23
Zabin et al. (1986a) Posttest only used for pregnancy rates in meta-analysis	School-linked program (university school of medicine); Baltimore	Pretest, posttest, follow-up over 2-year period/3	Sexual activity		
			481	–0.16	–0.08
			Contraceptive use		
			573	0.11	0.06
			Pregnancy rates		
			368	0.35	0.18
			46	–0.02	–0.01
Zabin et al. (1986b)	School-linked program (university school of medicine); Baltimore	Quasi-experimental; matched schools as comparison groups, pretest, posttest/3	Contraceptive use		
			2220	0.29	0.14
			333	0.11	0.05
			Sexual activity		
			716	–0.10	–0.05

*Due to missing data and lack of clarity in the reporting of statistical data within noted studies, partial data from studies were used in some meta-analyses. This resulted in the modification of the study design. Modifications are noted beside the authors' names.

†Some studies also measured additional outcomes.