

B U T D O E S I T W O R K ?

IMPROVING EVALUATIONS OF SEXUALITY EDUCATION



SEXUALITY INFORMATION AND EDUCATION COUNCIL OF THE UNITED STATES

BUT DOES IT WORK? IMPROVING EVALUATIONS OF SEXUALITY EDUCATION

Debra W. Haffner, M.P.H.
SIECUS President & CEO

Eva S. Goldfarb, Ph.D.
Assistant Professor, Montclair State University,
Upper Montclair, NJ

With support from the California Wellness Foundation, SIECUS convened a symposium of 15 of the nation's most prominent researchers in sexuality education and teenage pregnancy prevention in October 1996 to help improve evaluations of sexuality education programs.

They addressed such questions as:

- What is the status of current evaluation research on sexuality education? What are the gaps in the literature?
- Can the impact that sexuality education has on body image be evaluated? On self-esteem? On relationships? On adult sexual health?
- What methodologies can improve evaluations of sexuality education programs?
- What do program managers, evaluators, and funders need to know?

Unfortunately, evaluations of comprehensive sexuality education have tended to focus primarily on whether the programs have helped young people delay sexual activity and prevent unwanted pregnancy and disease. Evaluators have often defined program effectiveness as helping young people either postpone sexual intercourse or increase the use of contraception and condoms.

Other program goals of comprehensive sexuality education—such as helping young people develop an appreciation of their bodies or communicating effectively with peers and partners—are often overlooked in evaluations of sexuality education programs.

Comprehensive sexuality education has four goals:

- To provide young people with accurate information about human sexuality.
- To provide an opportunity for young people to question, explore, and assess their sexual attitudes.
- To help young people develop interpersonal skills, including communication, decision-making, assertiveness, and peer refusal skills, as well as the ability to create satisfying relationships.

- To help young people exercise responsibility regarding sexual relationships, including addressing abstinence, how to resist pressures to become prematurely involved in sexual behaviors, and encouraging the use of contraception and other sexual health measures.¹

The National Guidelines Task Force, convened by SIECUS to develop the *Guidelines for Comprehensive Sexuality Education: Kindergarten–12th Grade*, identified 36 life behaviors of a sexually healthy adult that are the desired results of a K–12 sexuality education program.

They include interacting with both genders in respectful and appropriate ways, viewing family as a valuable source of support, practicing effective decision-making, and expressing one's sexuality while respecting the rights of others. Important outcomes include using contraception, preventing sexual abuse, avoiding sexually transmitted diseases, and practicing health-promoting behaviors, *but they are not the only outcomes of a comprehensive program.* (See “Life Behaviors of a Sexually Healthy Adult” on page 5.)

RECENT EVALUATIONS

Recent reviews of evaluations of effective sexuality education, teenage pregnancy prevention, and HIV prevention programs have found that quality sexuality education programs:

- increase knowledge;
- clarify values;
- increase parent–child communication;
- help young people delay the initiation of sexual intercourse;
- increase the use of contraception and condoms;
- do not encourage young people to begin intercourse; and
- do not increase the frequency of sexual intercourse.²

These reviews also describe common characteristics of effective programs. Specifically, they:

- target specific behaviors;
- are based on a theoretical model for behavior change;

- provide information about the risks of unprotected sexual intercourse and how to reduce risk;
- provide students with an opportunity to practice skills and discuss situations that they find meaningful and realistic;
- address the influence of the media, peers, and culture on teenagers' sexual behaviors and decisions;
- develop and reinforce beliefs and values among students that support their decisions to be abstinent and/or to protect themselves; and
- include opportunities for students to practice communication and negotiation skills.³

FEW QUALITY EVALUATIONS

The symposium participants convened by SIECUS agreed that there are only a few well-designed, well-implemented, and well-funded evaluations of sexuality education. They also agreed that most published evaluations primarily determine the impact of programs based on three goals:

- delaying the onset of sexual intercourse;
- increasing contraceptive and condom use; and
- decreasing pregnancy and birth rates.

These goals are extremely difficult to attain given the limited nature (five to 12 sessions) of most programs.

There are few published evaluations concerning many of the desired outcomes of sexuality education such as appreciation of one's body; identifying and living according to one's values; developing and maintaining meaningful relationships; avoiding exploitative or manipulative behaviors; and engaging only in consensual relationships.

Generally, evaluators have not assessed the effect of such programs on helping young people achieve the life behaviors of a sexually healthy adult, and they have not developed instruments to measure these behaviors.

In addition, evaluations have usually only examined short-term effects. Most include immediate post-tests that do not allow for passage of time. Only a few examine the impact of the program beyond a year and a half. But, as the

SYMPOSIUM HIGHLIGHTS

The symposium held in October 1996 offered a unique opportunity for an exchange of ideas about the evaluation of sexuality education. Participants presented information about their own research, examined each other's studies, and explored new directions.

Although SIECUS did not ask participants to reach a consensus, certain important themes emerged.

LIMITATIONS OF EXISTING EVALUATIONS OF SEXUALITY EDUCATION PROGRAMS

- There have been too many weak evaluations of sexuality education and teenage pregnancy prevention programs.
- There are very few published evaluations of many of the desired outcomes of sexuality education (for example, appreciating one's body; identifying and living according to one's values; developing and maintaining meaningful relationships; avoiding exploitative or manipulative behaviors; communicating effectively with parents; and engaging only in consensual relationships).
- Sexuality education programs *alone* will *not* lower teenage pregnancy or birth rates or the incidence of STDs and HIV.

RECOMMENDATIONS FOR EVALUATIONS OF SEXUALITY EDUCATION PROGRAMS

- Doing no evaluation is better than doing a bad evaluation.
- Rigorous outcome evaluations are appropriate only after a program is successfully implemented for a period of years.
- Evaluations must be consistent with the expressed goals of the program and the course content.
- Evaluations of comprehensive sexuality education should go beyond measuring changes in whether young people are having intercourse or whether they are using a contraceptive method.
- Evaluations of *school-based sexuality education* should focus on changes in knowledge, attitudes, and skills. Be cautious about measuring outcomes outside the classroom.
- Simple programs should be evaluated by simple measures; complex outcome measures on behavior change should be reserved for multi-year, intensive strategies.
- There is a need for new instruments to measure sexual health objectives.
- Qualitative methods are an important supplement to quantitative methods.

sexuality educators at the symposium articulated, the young people may not experience results from the program until they start dating, leave home, or become adults.

Further, few program evaluations have been replicated. In those cases when replication has occurred, the results in another setting are often not as positive. As one symposium participant pointed out, "Successful programs may have more to do with individual teachers and their charisma, than with the actual program components." Conversely, one participant warned about the challenge of replicating a program with fidelity: "Is there really integrity in a replication from one site to another? My hunch is that 90 percent of the time, the only thing the programs have in common is the booklet the kids receive."

As a result of many of these issues, comprehensive, methodologically rigorous evaluations are extremely costly and time-consuming, and are, therefore, inaccessible to most programs.

UNIQUE CHALLENGES

Symposium participants agreed that comprehensive sexuality education presents unique challenges for program evaluations.

First, they acknowledged that changing human behavior is difficult and that simple educational efforts themselves have often met with limited success. They pointed out that most school-based courses are considered successful if they increase a student's knowledge and performance on standardized tests. Only health education efforts are held to a standard of behavior change outside of the classroom.

Second, they acknowledged that, unlike other academic programs, young people often learn about sexuality from a wide range of sources outside of school such as families, television, movies, advertising, peers, magazines, partners, church, and youth organizations. Yet evaluations tend to focus only on the impact of the school program: "Studies typically measured only the incremental effect of the intervention and not the cumulative effect of that intervention plus whatever sexuality education or reproductive health service the youth had previously or subsequently received. If the cumulative effect of all reproductive health education and services were measured, results would probably be stronger."⁴

Both youth and adults approach sexuality education with established attitudes and beliefs. People come into programs with their own previous knowledge, their own family values, their own cultural values, their own experiences, and their own fears. The reality is that students in any class have very different needs for sexuality information. Some need basic information about their bodies; some need to know how to handle a sexually abusive stepparent; some need to look at their personal values; and some need to know how to set sexual limits with their partner. In addition, some are not dating; some are abstinent from all sexual behaviors;

some are experimenting sexually; and some are having sexual intercourse. Most are heterosexual although some are bisexual, gay, lesbian or questioning their orientation.

Several participants commented on the fact that sexuality education affects people in different ways. In observing highly competent sexuality educators, a person senses that the young people in the class are excited, challenged, and learning. Yet, formal evaluations of these same programs have not always demonstrated positive results. The reason? As one participant noted, "For one person, one thing happens; for another person, something else happens. But if you measure only one behavior, and it was significant for only a few kids, it's lost and we don't capture it."

Participants said that evaluators need to better assess students' knowledge and attitudes at the baseline. What issues do they have coming into the program? Are they ready to learn what the program has to offer? How are they changed as a result of the program?

In addition, previous exposure to sexuality information complicates evaluation methodology. This may make it impossible to have true random experimental designs with random assignments to programs. Because people are exposed to both formal and informal "sexuality education" throughout their lives, there is probably no such thing as a true control group that has not been exposed to any information.

Third, classroom behaviors are often only proximal to behaviors in real life. For example, "if someone can negotiate condom use in a skill-building exercise in a classroom, do we really know whether they can do it in real life?" It is easier to assess whether a child can read or add a column of numbers both in school and outside of school than it is to know if they will practice effective decision-making or sexual limit-setting outside the classroom.

Fourth, sexuality education is one of the only school-based programs that measures behaviors outside of the classroom. Most school-based programs are simply evaluated by how young people score on tests. Although studies show that sexuality education increases knowledge, it is sometimes seen as not effective because it does not change sexual behaviors. In the words of one participant, "We don't measure the success of the math curriculum on whether people can and do balance their checkbooks."

Fifth, sexual behaviors involve more than the classroom participant. In most cases, only one member of the relationship has participated in the program; yet, that person is also expected to influence the behavior of their partner.

Sixth, asking young people about their sexual behaviors raises particular methodological concerns. Research that relies on self-report about sexual behaviors must take into account self-report bias and the problem of "social desirability" (giving answers that they believe the researcher wants to hear). At pre-test and post-test, participants may under-

LIFE BEHAVIORS OF A SEXUALLY HEALTHY ADULT

A sexually healthy adult will:

In human development

- Appreciate his or her own body.
- Seek further information about reproduction as needed.
- Affirm that human development includes sexual development that may or may not include reproduction or genital sexual experience.
- Interact with both genders in respectful and appropriate ways.
- Affirm his or her own sexual orientation and respect the sexual orientation of others.

In relationships

- View family as a valuable source of support.
- Express love and intimacy in appropriate ways.
- Develop and maintain meaningful relationships.
- Avoid exploitative or manipulative relationships.
- Make informed choices about family options and relationships.
- Exhibit skills that enhance personal relationships.
- Understand how cultural heritage affects ideas about family, interpersonal relationships, and ethics.

In personal skills

- Identify and live according to his or her values.
- Take responsibility for her or her own behavior.
- Practice effective decision-making.
- Communicate effectively with family, peers, and partners.

In sexual behavior

- Enjoy and express his or her sexuality throughout life.
- Express his or her sexuality in ways that are congruent with his or her values.
- Enjoy sexual feelings without necessarily acting on them.
- Discriminate between life-enhancing sexual behaviors and those that are harmful to self and/or others.

- Express his or her sexuality while respecting the rights of others.
- Seek new information to enhance his or her sexuality.
- Engage in sexual relationships that are consensual, non-exploitative, honest, pleasurable, and protected against disease and unintended pregnancy.

In sexual health

- Use contraception effectively to avoid unintended pregnancy.
- Prevent sexual abuse.
- Act consistent with his or her own values in dealing with an unintended pregnancy.
- Seek early prenatal care.
- Avoid contracting or transmitting an STD, including HIV.
- Practice health-promoting behaviors, such as regular check-ups, breast and testicular self-exam, and early identification of potential problems.

In society and culture

- Demonstrate respect for people with different sexual values.
- Exercise democratic responsibility to influence legislation dealing with sexual issues.
- Assess the impact of family, cultural, religious, media, and societal messages on his or her thoughts, feelings, values, and behaviors related to sexuality.
- Promote the rights of all people to accurate sexuality information.
- Avoid behaviors that exhibit prejudice and bigotry.
- Reject stereotypes about the sexuality of diverse populations.
- Educate others about sexuality.

Resource: *Guidelines for Comprehensive Sexuality Education: Kindergarten–12th Grade, 2nd Edition* (New York: SIECUS, 1996).

report their involvement in risky or unhealthy behaviors, and, thus, reduce the measured impact of the program.

Seventh, there are often policy barriers to high quality evaluations of sexuality education programs. Evaluations of sexuality education may be too controversial in some communities to enable a thorough study. Legislation within many states, as well as the proposed Family Privacy Act at the federal level, limit the types of data that can be collected on sexual behavior. In some states, researchers must obtain parental consent before they can ask a young person about their sexual attitudes and behaviors.

Eighth, the group cautioned about treating sexuality education as synonymous with teenage pregnancy or HIV-prevention efforts. Teenage childbearing is affected by many social and economic factors such as poverty, racism, sexism, job opportunities, past history of sexual abuse, family stability, school failure, and risk-taking behaviors. Sexuality education is a necessary component of teenage pregnancy prevention efforts, but it cannot solve the problem alone.

One participant noted, “I say to communities, do you think it is possible to prevent pregnancies if our young people don’t know where babies come from?...We ought to be asking communities, ‘Do you have the sexuality education in place so that you’re sure all young people are sexually literate?’ But communities need to understand that the goals of teenage pregnancy prevention are very specific and not solved by short-term school-based interventions.”

Another said that “simply addressing sexual beliefs, attitudes, and skills—and even improving access to contraception—will not address many of the factors leading to teenage childbearing, may not significantly change young people’s motivation to avoid childbearing, and are unlikely to significantly reduce long-term sexual risk-taking.”

EVALUATION METHODOLOGY

Rigorous evaluations share common criteria. Quality evaluations share similar characteristics. They

- evaluate a sufficient number of representative programs;
- use random assignment;
- include a sufficiently large sample size;
- conduct long-term follow-up;
- measure behavior rather than attitudes or beliefs;
- conduct proper statistical analyses;
- publish both positive and negative results;
- replicate studies of successful programs; and
- use independent evaluators.⁵

It is not easy or always possible, however, to incorporate all of these criteria into every study design because of the significant time and money investments.

Symposium participants cautioned that not all programs require this level of evaluation. They therefore made some recommendations for both evaluators and program personnel that follow.

- Recognize the difference between evaluations that are designed to provide input into program implementation and those that are designed to measure impact as well as for publication in the professional literature.
- Remember that programs need different types of evaluations at different stages. In the words of one participant, “Evaluations need to be responsive to the different stages of program development and maturation.”
- Focus modest evaluations on a few key outcomes. One participant advised, “Ask yourself, ‘What are the three things I should look at to determine whether this program is effective or not?’”
- Remember that although large numbers are necessary for statistical significance, many programs do not warrant this level of evaluation. According to one participant, “Not everything is worth a 2,000 child random-assignment design.” The participants agreed that no evaluation is better than a bad evaluation.

RESEARCH METHODS

Symposium participants did not extensively address basic research methods. The “Recommended Resources” listed in this article will, however, provide extensive references for people new to program evaluation. (See page 12.) In addition, “Methods of Measurement” provides a comparison of some of the most frequently used measurements that evaluators might consider in developing designs. (See page 13.)

Evaluation designs can be divided into two categories: *quantitative* and *qualitative*. By using a quantitative design, the evaluator can answer questions that have to do with “how well” a program works, “to what extent” a program achieves its goals and objectives, “how much” learning and/or behavior change takes place, and “to what degree” to attribute measured outcomes to the evaluated program.

Quantitative evaluations are often very powerful and well-controlled. They cannot, however, answer questions about “why” a program is having some effect or “how” the program has the impact that it does. If a program is found effective (or ineffective) in reaching its objectives, is it because of one specific aspect such as the teacher, the setting, the participants, the environment, or some combination of these factors? For these types of questions, qualitative evaluation models may prove more useful.

Participants at the symposium urged that evaluations of comprehensive sexuality education programs combine both quantitative and qualitative approaches. Although these designs are often seen as opposites, evaluations are, in fact, often made much stronger by combining methods from both approaches—a strategy known as *triangulation*.

Qualitative methods are often extremely useful in helping to form the questions for an evaluation, in describing processes in particular programs, and in elucidating findings from traditional quantitative methodologies. In many cases, the emphasis of qualitative evaluation is on producing information that is useful for the further development of programs. Such evaluations typically do not lead to conclusions about whether a program was successful or unsuccessful, but, instead, produce detailed descriptions of how participants experience a program and its operation. From these descriptions, program directors can make decisions about future directions. Data collected from these evaluations are often given to program staff as the process is ongoing rather than after the data is collected.⁶

There are three usual objectives of qualitative evaluations.

- They produce as complete a description as possible of the components of the program, particularly the interactions among program staff and students.
- They produce a complete description of the context within which the program operates. What external forces affect the development and direction of the program? What beliefs, skills, and knowledge do participants bring to the program? What other social and political factors have an impact on the program?
- They create a picture of how staff and participants view and come to understand the program. The evaluator attempts to learn how it feels to be a teacher or student in the program and to describe what kinds of changes they go through in the process.⁷

LIFE BEHAVIORS OF A SEXUALLY HEALTHY ADULT

Participants at the symposium spent considerable time on how to best evaluate progress toward achieving the desired life behaviors of a sexually healthy adult. The discussion included using established instruments and data collection techniques to measure variables specific to sexuality education. Participants also discussed the challenge for researchers to develop new instruments to measure variables not yet studied.

Discussions about measuring the impact of programs yielded the following general suggestions:

- Include a broad qualitative component. This could include letters from former students articulating specific

ways in which a program had an impact, in-depth interviews with students with open-ended questions that would allow them to emphasize what they found important, and discussions with teachers about what they felt they had accomplished.

- Utilize a portfolio-style assessment that combines different documents (such as students' journals) through which they can discuss their recent real-life experiences. By using students' own words, researchers can discover what was significant to them.
- Observe students to measure attitudes, behaviors, and comfort levels. For example, videotape classroom interactions at the beginning and end of a program looking for changes in the ways students interact with each another with regard, for example, to having respect for both genders, or comfort discussing sexuality-related issues.
- Ask students to monitor their own behaviors both in and out of the classroom, and to report on them in class or homework. This could measure progress on certain skills, and serve as a teaching tool.
- Use focus groups at the end of a program to ask "How has this program affected you?" "What have been the most important aspects of the program for you personally?" "Did this program change the way you think about your sexuality?" As one meeting participant suggested, a person can learn why the program worked, with whom it worked the best, and what elements connected with certain kinds of young people.
- Look at long-term or delayed results. Talk with students five years after they have participated in a program. Ask them what impact, if any, their sexuality education had on their lives.

The group also spent time thinking about possibilities for measuring specific life behaviors, including adapting current instruments and developing new measurement approaches. The following ideas relate to some of the behaviors that might be measured. Many of the suggestions are, however, applicable to other variables.

Appreciation of one's own body. There are a number of well-tested scales available to measure attitudes about one's self and one's body image. Most of them focus on post-pubertal body image.⁸ Because body image is an important variable that is related to other aspects of self-image, programs may want to focus on body image issues earlier in young people's lives. The development of scales that measure pre-pubertal body image (in males and females) is an important area for instrument development.

Communicating effectively with parents, peers, and partners. There are many instruments to measure general communication skills with children, with adults, and across age groups. What is largely missing, however, are instruments that measure specific communication skills about sexuality issues. While some communication skills are applicable to many different types of situations, there may be specific skills related to communicating around sexual behaviors (for example, communicating about condom use or establishing sexual limits). In creating these instruments, two considerations are paramount: First, how does

one measure the quality and content, as well as the quantity of communication? Second, when measuring communication skills, it is important to assess the communication from the perspectives of all of the people involved in the interaction.

Practicing effective decision-making. A suggested technique for measuring this skill is to provide scenarios to students that require them to decide what they would do in a certain situation, how they would respond, and what alternatives they would consider.

SYMPOSIUM PARTICIPANTS

Claire Brindis, Dr.P.H.

Center for Reproductive Health Policy Research
University of California/San Francisco
San Francisco, CA

Janet Collins, Ph.D.

U.S. Centers for Disease Control and Prevention
Atlanta, GA

Jacqueline E. Darroch, Ph.D.

The Alan Guttmacher Institute
New York, NY

Eva S. Goldfarb, Ph.D.

Montclair State University
Upper Montclair, NJ

Douglas Kirby, Ph.D.

ETR Associates
Santa Cruz, CA

Barbara Marquez

California Department of Health Services
Office of Family Planning
Sacramento, CA

Dennis McBride, Ph.D.

Washington Institute
University of Washington
Tacoma, WA

Guy Parcel, Ph.D.

Center for Health Promotion
University of Texas Health Service Center
Houston, TX

Susan Philliber, Ph.D.

Philliber Research Associates
Accord, NY

Michael Resnick, Ph.D.

National Teen Pregnancy Prevention Research Center
University of Minnesota
Minneapolis, MN

Mark Roosa, Ph.D.

Arizona State University
Tempe, AZ

Sharon Thompson

Author
New York, NY

Mohammad R. Torabi, Ph.D., CHES

Indiana University
Bloomington, IN

Bonnie Trudell, Ph.D.

Department of Public Instruction
Madison, WI

Brian Wilcox, Ph.D.

University of Nebraska
Lincoln, NE

William Yarber, H.S.D.

Indiana University
Bloomington, IN

CALIFORNIA WELLNESS FOUNDATION

Mercedes Siordia

Gayle Wilson

SIECUS STAFF

Debra W. Haffner, M.P.H.

Leslie Kantor, M.P.H.

Identifying and living according to one's values. One suggestion for measuring this behavior is to have young people identify their three most important values around sexuality and then to ask them, "How does this influence your life?" Evaluators could use semantic differential scales that present two polar opposites along a continuum in relationship to some value to discover students' values. For example, a scale could have "sexually ethical" at one end and "sexually unethical" at the other. A student would choose a place along the continuum that best represented his or her values. The evaluator could then follow up with a question such as "How certain are you that this is your value?"

Interacting with both genders in respectful and appropriate ways. One problem with measuring this and certain others is in defining the related concepts in a measurable way, an action known as *operationalization*. Meeting participants acknowledged that, in some cases, it is easier to measure the absence of certain negative behaviors. For example, an individual could observe interactions of participants and define "respectful and appropriate" as the absence of physical violence, verbal abuse, exploitation, bullying, lying, and tricking. While these are easier to define, there are potential problems in defining a concept by what it is not.

One way of measuring positive indicators of relationships among people is to use a variety of "social distance" scales that examine the degree to which people are comfortable being close to other groups of people.⁹ They could, for example, measure a person's comfort with people of the other gender as well as people of different sexual orientations, people with different sexual values, and people living with AIDS.

Participants warned, however, that some of the life behaviors do not easily lend themselves to evaluation. They also cautioned that some life behaviors are no easier to achieve than preventing teenage pregnancy. In the words of one participant, "Which is harder, preventing pregnancy or promoting gender equality?" One asked, "Should we really hold a 10-hour program to the standard of seeing pre-/post-test change in gender relationships?"

RECOMMENDATIONS FOR IMPROVING EVALUATIONS

Program directors, funding agencies, and evaluation experts are interested in evaluations for different reasons. These are the words of symposium participants:

An Evaluation Expert "The scientist in me wants to know, 'Does a program work?' And if so, 'Why?' And if not, 'Why not?'"

A Program Director "Program directors are interested in evaluations to improve the performance of their programs,

to see whether they are really reaching the people they want to reach and whether the services they want to deliver are delivered to the right numbers of the right people effectively.... In other words, 'Is my program doing a good job and is it doing what I'm hoping it's doing?'"

A Foundation Executive "Is our program support really making a difference in the lives of the people we care about?"

The symposium participants offered suggestions for improving evaluations of comprehensive sexuality education. The recommendations are targeted at three main groups: program directors, evaluators, and funders. Many apply to all three.

PROGRAM DIRECTORS

Recommendations for program directors include:

- Be realistic about what your program can accomplish.
- Develop programs and evaluations based on well-tested theory, such as social learning theory.
- Do not oversell your program or make claims that are not attainable; don't set yourself up for failure.
- Both the program and the evaluation you design must be in line with available resources. If your budget is small, plan a small-scale program and simple evaluation.
- Plan the program evaluation at the beginning. Build the evaluation team into the program planning process. Consider seeking funding in partnership with an evaluation team.
- Obtain a qualified evaluator. Use an external evaluator:
 - with experience in, and understanding of, the policies, issues, and politics of sexuality education;
 - with specific, practical experience in evaluation;
 - with an understanding of the evaluation process and methods, including quantitative and qualitative measures;
- Ask the proposed evaluator for samples of her or his work and check references.
- Make sure that the person you contract to conduct the evaluation is actually the person who does it.
- Stay involved through the evaluation process. Make your interest and concerns known.
- Involve the evaluator early. Let him or her review the content and implementation of the program from the start.
- Establish a partnership between program personnel and evaluation teams. Make sure that both are involved throughout the process.

- Have a plan up front for how you will use the evaluation results. This can help guide the evaluation team in deciding which questions to ask.
- Be aware of the political issues surrounding the evaluation of your program.
- Do not evaluate program outcomes too soon; give it time to become established and to have an impact. A program may have to be in place for several years and go through several adjustments before its true impact is measurable.
- Know how you are going to use any data you collect; only collect data that you really want. "If it's only nice to know, it's gotta go."
- Settle issues of potential conflict with evaluators. For example, who will own the data after it is collected and analyzed? In what sequence will the authors' names appear on any resulting publications? Who has final say on what and how findings are reported?
- Consider conducting your own process evaluation as you go along to help fine tune your program.

EVALUATORS

Recommendations for evaluators include:

- Address the broader goals for sexuality education. Pay attention to the life behaviors of a sexually healthy adult.
- Find ways to measure mediating variables that may be related to outcomes. For example, factors such as self-esteem, self-confidence, and the ability to behave according to one's own values may serve as the mediating factors between where participants are and where a program hopes to take them.
- Identify and report small effects that may apply only to a few people.
- Identify the people most directly involved in the program and involve them in the process.
- Evaluate the content and methodologies of a program before measuring student outcomes to make sure there is a match between the program's content and methodologies and its stated objectives.
- Monitor the program implementation to determine if participants are receiving what they are supposed to receive.
- Make certain you are knowledgeable about sexuality education.
- Consider how the personalities of the people in a program can affect its implementation. (Ethnographic methodologies are particularly well suited to such questions.)
- Do not assume the relevant questions; let them emanate from practice. Engage people involved in the program—including participants themselves—to formulate questions. This increases the validity of the evaluation and increases the chances for utilizing the information.
- Recognize that every person has probably had some exposure to some sexuality education or information before coming into a particular program. Try to measure this prior exposure. Obtain accurate baseline measurements.
- Consider measuring the volume and intensity or dosage effects within programs.
- Broadly disseminate both positive and negative findings. Consider a variety of formats for providing information in addition to peer review journals.
- Develop new instruments to measure the various variables related to the behaviors of a sexually healthy adult.
- Provide feedback to the institution and personnel involved with the program throughout the process. Early formative evaluation data is often useful in decision-making related to the program.
- Disseminate final findings to the involved institution. Give presentations, conduct in-service activities, give feedback reports, make presentations to school boards, and meet with teachers.
- Include teachers' perspectives in the evaluation. Find out what they think they are doing and what they would like to do in relation to the intervention.
- Use statistical analysis that is appropriate for the program.
- Measure changes in knowledge immediately; measure changes in attitudes at least three to six months after a program; measure changes in behavior at least 12 to 18 months after the program.
- Develop cultural competency. Make certain you recognize and take into account how the group you are evaluating understands sexuality issues.

FUNDERS

Recommendations for funders include:

- Have realistic expectations for program outcomes in relation to costs, funding levels, available time. Don't expect tremendous outcomes from a program that is funded for a brief time period.
- Consider contracting with an independent organization to evaluate the programs you support.
- Provide adequate resources for appropriate evaluations.

- Do not require a full-impact evaluation from every program. Sometimes a process evaluation to help improve a program is appropriate.
- Use evaluation findings to guide funding decisions but don't rely solely on a single evaluation to decide to continue supporting a program.
- If an evaluation finds that a program doesn't seem to work, it could be for a variety of reasons. Don't give up a program just because an evaluation doesn't show expected outcomes.
- Ask for line item budgets for evaluation components in grant applications.
- Accept process measures—variables related to larger behavioral outcomes but sometimes easier to measure—for some types of programs. “How many sessions did you run? How many people received your information?”
- Be realistic about the claims you make to your board of trustees or others about the programs you fund.
- Change focus from funding one program that “does it all” to several smaller programs and different groups of people involved in different aspects of the field.
- Provide technical assistance and/or support over the long-term to program personnel and evaluators to learn how to disseminate their findings and how to negotiate within the political arena that surrounds their work.
- Support training in evaluation for your grantees. Build the capacity for evaluation into the community organizations themselves.
- Provide support for training more evaluation experts who are from communities of color.

CONCLUSION

SIECUS urges program directors, evaluation experts, and foundations to assess how they can better evaluate the impact of sexuality education programs.

Such evaluations must move beyond simplistic “did you like it?” post-test questionnaires as well as more sophisticated

designs that measure only the impact on the incidence of intercourse and contraceptive use.

Practitioners have a critical need to understand what makes a program effective (and for which students) and how to improve their education programs. The symposium and this report offer an important first step in this effort.

Editor's Note: *SIECUS is grateful to the California Wellness Foundation for its generous support of the symposium and this report.*

REFERENCES

1. National Guidelines Task Force (1996), *Guidelines for Comprehensive Sexuality Education: Kindergarten–12th Grade*, 2nd Edition, Sexuality Information and Education Council of the United States.
2. A. Grunseit and S. Kippax (1993), “Effects of Sex Education on Young People's Sexual Behaviour,” unpublished review commissioned by the Global Programme on AIDS, World Health Organization, July 1993; K. A. Moore, et. al. (1995), *Adolescent Pregnancy Prevention Programs: Interventions and Evaluations*, Child Trends, Inc., Washington, DC; J. J. Frost and J. D. Forrest (1995), “Understanding the Impact of Effective Teenage Pregnancy Prevention Programs,” *Family Planning Perspectives* 25(5): 188-96; and D. Kirby et. al., “School-Based Programs to Reduce Sexual Risk Behaviors: A Review of Effectiveness,” *Public Health Reports*, 109(3), 1994, pp. 339-60.
3. D. Kirby et. al., “School-Based Programs to Reduce Sexual Risk Behaviors.”
4. D. Kirby, *No Easy Answers: Research Findings on Programs to Reduce Teen Pregnancy* (Washington, DC: The National Campaign to Prevent Teen Pregnancy, 1997).
5. Ibid., p. 3.
6. J. R. Schuerman, *Research and Evaluation in the Human Services* (New York: Macmillan, 1983), pp. 113-14.
7. Ibid., p. 114.
8. M. Story, “Body Attitudes Questionnaire,” in C. Davis, W. Yarber, and S. Davis, editors, *Sexuality-Related Measures: A Compendium* (Syracuse, NY: Syracuse University, 1987), p. 40.
9. E. R. Babbie, *Survey Research Methods* (Belmont, CA: Wadsworth Publishing Co., 1973).

RECOMMENDED RESOURCES

PUBLISHERS

Sage Publications

P.O. Box 5084
Thousand Oaks, CA 91359-9924
Phone: 805/499-9774
Fax: 805/499-0871
Web site: <http://www.sagepub.com>
E-mail: order@sagepub.com

- **Evaluation: A Systematic Approach** (1993)
P. Rossi and H. Freeman. \$44
- **Foundations of Program Evaluation** (1991)
W.T. Shadish, T.D. Cook and L. Leviton
\$65, hardcover; \$31.50, paperback
- **Handbook of Adolescent Sexuality and Pregnancy:
Research and Evaluation Instruments** (1993) \$62
- **Program Evaluation Kit** (1987) (9 volumes) \$95
- **Publications Catalog: Research Methods and
Evaluation** (1996) Free
- **The Survey Kit** (1995) (9 volumes) \$132.55

Education Program Associates

1 West Campbell Avenue—Building D
Campbell, CA 95008
Phone: 408/374-3720
Fax: 408/374-7385

- **Evaluating Your Information
and Education Project** (1988)
C.D. Brindis and P.B. Reyes. \$25
- **Evaluation Guidebook for Family Planning
Information and Education Projects**
C.D. Brindis, C.C. Korenbrot, P.L. Brown. \$20
- **Strategies and Evaluations in Programs
for Pregnant and Parenting Teens:
Lessons to Date and Future Directions**
S. Philliber and C.D. Brindis. \$20

Sociometrics Corporation

170 State Street, Suite 260
Los Altos, CA 94022-2812
Phone: 800/846-DISK
Fax: 415/949-3299
Web site: <http://www.socio.com>
E-mail: pasha@socio.com

- **Assessing Your Community's Needs and Assets:
A Collaborative Approach to Adolescent
Pregnancy Prevention**
C. Brindis, J.J. Card, S. Niego, and J. L. Peterson. \$20
- **Evaluating and Monitoring Programs for
Pregnant and Parenting Teens** (1988) \$20
- **Evaluating Programs Aimed at Preventing
Teenage Pregnancies** (1988) \$20
- **Local Evaluator Consultant Network** (1996) \$6
- **Prevention Minimum Evaluation Data Set
(PMEDS)** (1996) \$20 printed version only,
\$30 with diskette

JOURNALS

Sage Publications

(See contact information under *Publishers*.)

- **Evaluation and the Health Professions**
Quarterly; Institutional rate: \$198; Individual rate: \$62
- **Evaluation Review:
A Journal of Applied Social Research**
Bimonthly; Institutional rate: \$255; Individual rate: \$78

Heldref Publications

1319 18th Street, N.W.
Washington, DC 20036
Phone: 800/365-9753
Fax: 202/296-5149
E-mail: jer@heldref.org

- **Journal of Educational Research**
Bimonthly; Institutional rate: \$80; Individual rate: \$40;
Single copy: \$13.50

METHODS OF MEASUREMENT

Method	Description	Research Design	Advantages	Disadvantages
Questionnaire	A paper and pencil method for obtaining responses to statements or questions by using a form on which participants provide opinions or factual information.	Quantitative	<ul style="list-style-type: none"> • Relatively inexpensive, quick way to collect large amounts of data from large samples in short amount of time; • Convenient for respondents to complete; • Anonymity can result in more honest responses; • Questionnaires are available; • Well suited for answering questions related to "What?" "Where?" and "How many?" 	<ul style="list-style-type: none"> • Limited ability to know if one is actually measuring what one intends to measure; • Limited ability to discover measurement errors; • Question length and breadth are limited; • No opportunity to probe or provide clarification; • Relies on participants' ability to recall behavior, events; • Limited capability to measure different kinds of outcomes; • Must rely on self-report; • Not well suited to answering questions related to "How?" and "Why?"; • Difficult with low-literacy groups.
One-to-One Interview	An interaction between two people in which information is gathered relative to respondent's knowledge, thoughts and feelings about different topics.	Quantitative/ Qualitative	<ul style="list-style-type: none"> • Allows greater depth than questionnaire; • Data is deeper, richer, has more context; • Interviewer can establish rapport with respondent; • Interviewer can clarify questions; • Good method for working with low-literacy respondents; • Higher response and completion rates; • Allows for observation of nonverbal gestures. 	<ul style="list-style-type: none"> • Requires a lot of time and personnel; • Requires highly trained, skilled interviewers; • Limited number of people can be included; • Is open to interviewer's bias; • Prone to respondents giving answers they believe are "expected" (social desirability); • No anonymity; • Potential invasiveness with personal questions.
Unstructured	Totally free response pattern; allows respondent to express ideas in own way and time.	Qualitative	<ul style="list-style-type: none"> • Can elicit personal information; • Can gather relevant unanticipated data; • Interviewer can probe for more information. 	<ul style="list-style-type: none"> • Requires great skill on part of interviewer; • More prone to bias in response interpretation; • Data are time-consuming to analyze.
Semistructured	Limited free response, built around a set of basic questions from which interviewer may branch off.	Qualitative/ Quantitative	<ul style="list-style-type: none"> • Combines efficiency of structured interview with ability to probe and investigate interesting responses. 	<ul style="list-style-type: none"> • Cannot do true exploratory research; • Predetermined questions limit ability to probe further.

Table continues on next page.

Method	Description	Research Design	Advantages	Disadvantages
Structured	Predetermined questions, often with structured responses.	Quantitative	<ul style="list-style-type: none"> • Easy to administer; • Does not require as much training of interviewer. 	<ul style="list-style-type: none"> • Less ability to probe for additional information; • Unable to clarify ambiguous responses.
Focus Group	Interviews with groups of people (anywhere from four to 12) selected because they share certain characteristics relevant to the question of study. Interviewer encourages discussion and expression of differing opinions and viewpoints.	Qualitative	<ul style="list-style-type: none"> • Studies participants in natural, real-life atmosphere; • Allows for exploration of unanticipated issues as they are discussed; • Can increase sample size in qualitative evaluation; • Can save time and money; • Can stimulate new ideas among participants; • Can gain additional information from observation of group process; • Can promote greater spontaneity and candor. 	<ul style="list-style-type: none"> • Interviewer has less control than in a one-to-one interview; • Data is sometimes difficult to analyze; • Must consider context of comments; • Requires highly trained observer-moderators; • Cannot isolate one individual's train of thought throughout.
Phone Interview	One-to-one conversation over the phone.	Qualitative/ Quantitative	<ul style="list-style-type: none"> • Potentially lower cost; • Anonymity may promote greater candor. 	<ul style="list-style-type: none"> • Not everyone has a phone; • Unlisted numbers may present sampling bias; • No opportunity to observe nonverbal gestures.
Participant Observation	Measuring behaviors, interactions, processes by directly watching participants.	Qualitative/ Quantitative	<ul style="list-style-type: none"> • Spontaneous quality of data that can be gathered; • Can code behaviors in a natural setting such as a lunch room or hallway; • Can provide a check against distorted perceptions of participants; • Works well with a homogeneous group; • Good technique in combination with other methods; • Well suited for study of body language (kinesics) and study of people's use of personal space and its relationship to culture (proxemics). 	<ul style="list-style-type: none"> • Quantification and summary of data is difficult; • Recording of behaviors and events may have to be made from memory; • Difficult to maintain objectivity; • Very time-consuming and expensive; • Requires a highly trained observer.
Participant As Observer	The evaluator's role as observer is known to the group being studied and is secondary to his or her role as participant.	Qualitative/ Quantitative	<ul style="list-style-type: none"> • Evaluator retains benefits of participant without ethical issues at stake. 	<ul style="list-style-type: none"> • Difficult to maintain two distinct roles; • Other participants may resent observer role; • Observer's presence can change nature of the interactions being observed.

Method	Description	Research Design	Advantages	Disadvantages
Observer As Participant	Evaluator's observer role is known and his or her primary role is to assess the program.	Qualitative/ Quantitative	<ul style="list-style-type: none"> Evaluator can be more focused on observation role while still maintaining connection to other participants. 	<ul style="list-style-type: none"> Evaluator is clearly an outsider; Observer's presence can change nature of the interactions being observed.
Complete Observer	The evaluator has no formal role as participant; is a silent observer; may also be hidden from the group or in a completely public setting where his or her presence is unnoticed and unobtrusive.	Qualitative/ Quantitative	<ul style="list-style-type: none"> More objective observations possible; Evaluator is not distracted by participant role; Evaluator's observations do not interfere in any way with the group's process if his or her presence is hidden. 	<ul style="list-style-type: none"> If evaluator's presence is known, it can inhibit or change interactions of participants; If evaluator's presence is hidden, it raises ethical questions.
Document Analysis	Unobtrusive measure using analysis of diaries, logs, letters, and formal policy statements to learn about the values and beliefs of participants in a setting or group. Can also include class reviews, letters to teachers, letters from parents, and letters from former students to learn about the processes involved in a program and what may be having an impact.	Quantitative/ Qualitative	<ul style="list-style-type: none"> Diaries reduce problems of memory relating to when, where, with whom; Provides access to thoughts and feelings that may not otherwise be accessible; Can be less threatening to participants; Evaluator can collect and analyze data on own schedule; Relatively inexpensive. 	<ul style="list-style-type: none"> Quality of data varies between subjects; Diaries may cause change in subjects' behaviors; Not well suited for low-literacy groups; Can be very selective data; No opportunities for clarification of data.
Archival Data	Analysis of archival data from a society, community or organization. Can include birth rates, census data, contraceptive purchase data, and number of visits to hospitals for STDs.	Quantitative	<ul style="list-style-type: none"> More accurate than self-report. 	<ul style="list-style-type: none"> Not all data is available or fully reported; Difficult to match data geographically or individually.
Historical Data	Analysis of historical data is a method of discovering, from records and personal accounts, what happened in the past. It is especially useful for establishing a baseline or background of a program or of participants in a program prior to measuring outcomes.	Quantitative/ Qualitative	<ul style="list-style-type: none"> Baseline data can help with interpretation of outcome findings; Can help answer questions about why a program is or is not successful in meeting its goals; Provides a picture of the broader context within which a program is operating. 	<ul style="list-style-type: none"> Can be difficult to obtain data; Relies on data that may be incomplete, missing, or inaccurate; May rely on participant's selective memory of events and behaviors; Difficult to verify accuracy.

M i s s i o n

SIECUS affirms that sexuality is a natural and healthy part of living. SIECUS develops, collects and disseminates information, promotes comprehensive education about sexuality, and advocates the right of individuals to make responsible sexual choices.



1 3 0 W E S T 4 2 N D S T R E E T , S U I T E 3 5 0 , N E W Y O R K , N Y 1 0 0 3 6