NSF

EVALUATION HIGHLIGHTS

A Report on the National Science Foundation's Efforts to Assess the Effectiveness of Its Education Programs

AUGUST 1996

Directorate for Education and Human Resources

Division of Research, Evaluation and Communication

An REC Report on Evaluation

August 1, 1996

Dr. Luther S. Williams
Assistant Director
Directorate for Education and Human Resources

Dear Dr. Williams:

I am pleased to transmit *NSF Evaluation Highlights: A Report on the National Science Foundation's Efforts to Assess the Effectiveness of Its Education Program.* This report is an edited compilation of evaluation reports produced since 1992, when you created an Evaluation unit (now housed in the Division of Research, Evaluation and Communication, or REC) within the Directorate. It was prepared with the assistance of Westat, Inc.

This report reflects how EHR program evaluations inform what we know about the impacts of NSF's education investments; how the results of these independent evaluations provide information relevant to the shaping and management of the EHR portfolio; and how the ongoing schedule of evaluations demonstrates staff and program accountability to the NSF leadership, the National Science Board, the Congress, and the American public.

Readers will draw their own conclusions about the utility of this evaluation work. What cannot be denied is NSF's dedication to demonstrating how collections of projects in localities throughout the Nation — subsumed under EHR program names, supported through merit-based competitions, and systematically monitored postaward — allow us to gauge progress in improving science, mathematics, engineering, and technology education at all levels of the system, kindergarten to participation in the workforce. We are not merely claiming to make a difference. Program evaluations try to measure the magnitude and pace of that difference in teaching and learning.

It is therefore with pride that I submit this report. It is a testament to your decision to require program evaluation as an integral part of EHR's "business." I hope you find *NSF Evaluation Highlights* a useful "horse" in the EHR accountability stable.

Sincerely yours,

Daryl E. Chubin
Division Director for Research, Evaluation and Communication

FOREWORD

Since 1992, the National Science Foundation's Directorate for Education and Human Resources (EHR) has been engaged in an effort to evaluate all of its science and mathematics education programs. The purpose of the evaluation initiative—which is orchestrated by the Directorate's Division of Research, Evaluation and Communication (REC)—is essentially twofold: to provide EHR officials with information that will help them manage more effectively the approximately 30 programs in the Directorate's portfolio; and to report to Congress and the public on the effectiveness of the Foundation's science and mathematics education programs.

To date, the evaluation effort has yielded more than a dozen reports, including several extensive evaluations that have been completed and briefer studies focusing on the ongoing impact of programs whose evaluations are in progress. EHR is observing a schedule that calls for full evaluations of all of its science and mathematics education programs to be accomplished within the next few years.

The first section of this four-part document (The Value of Evaluation) presents an overview of REC's evaluation efforts; the second (Highlights) summarizes the reports completed thus far; the third (Evaluation-Related Activities) describes REC's participation in evaluation-oriented projects that extend beyond conducting program evaluations for EHR; and the concluding section (Future Evaluation Efforts) discusses the Directorate's future objectives and schedule for fulfilling its evaluation objectives.

TABLE OF CONTENTS

THE VALUE OF EVALUATION

HIGHLIGHTS

EVALUATION-RELATED ACTIVITIES

FUTURE EVALUATION EFFORTS

Driven by the perception that primary and secondary science and mathematics education is in need of significant improvement throughout the United States, the National Science Foundation (NSF) in 1991 launched a major effort called the Statewide Systemic Initiatives Program (SSI).

This program, initiated by the Foundation's Directorate for Education and Human Resources (EHR), was created on the premise that positive reform in K-12 science and mathematics education will effectively be achieved if pursued by means of a systemic effort, coordinated nationwide on the state level rather than piecemeal. The aim of systemic reform is for all states ultimately to participate in a comprehensive endeavor to elevate teaching standards, to enrich the instructional materials that are offered to students, and to refine the methods employed in measuring students' grasp of science- and mathematics-related subjects. Central to the systemic reform concept is that all students should benefit, regardless of gender, ethnic, linguistic, or disability status.

> The ultimate aim of the evaluation initiative is to make sure EHR is fulfilling its mission to improve the Nation's science, mathematics, engineering, and technology education.

To bolster this ambitious undertaking, EHR, through its SSI program, has over the past 5 years provided up to $2 million annually to individual states offering substantive plans for implementing systemic reform.

Is the EHR initiative paying off? Is science and mathematics education in the Nation's primary and secondary schools getting better? If so, how much K-12 improvement is attributable to SSI? And how might EHR enhance the program to hasten achievement of its conceptual goals? To explore such issues and to monitor the program as it evolves, EHR initiated a large-scale evaluation of SSI in 1992.

Although the evaluation, conducted by outside contractors, is due to be completed in 1997, it already has yielded three extensive reports and a wealth of useful insights. Judging from the evaluators' reports, SSI is indeed succeeding overall. Most participating states are developing new and more effective science and mathematics curricula; they are working toward the clarification of student achievement criteria; more effort is being made to elevate the quality of science and mathematics teaching; and more state and local funding is being attracted to complement NSF's financial support of systemic reform.

The evaluations are primarily guided by three fundamental questions: Is the program under study achieving its goals? Is it making an impact? And are there ways in which the program can be improved?

Not so gratifying, however, are revelations—also stemming from the evaluations—that the SSI program, with all of its virtues, is lagging in some respects. Some states, for example, have been slow to develop strategies in response to the SSI challenge, while others have failed to galvanize public support for improvement in science and mathematics education.

Today, SSI continues as a major component of the EHR program portfolio—its objectives, expectations, financial viability, and other matters constantly undergoing the scrutiny of the evaluation team. This evaluation, however, is but one example of an unprecedented NSF initiative currently underway to subject all of the EHR directorate's programs to sound, rigorous, and thorough examination by outside analysts.

The evaluations are primarily guided by three fundamental questions: Is the program under study achieving its goals? Is it making an impact? And are there ways in which the program can be improved? The ultimate aim of the evaluation initiative is to make sure EHR is fulfilling its mission to improve the Nation's science, mathematics, engineering, and technology education. Broadly speaking, EHR's responsibility entails the development and administering of programs designed to ensure that high-quality science and mathematics education is available to every child in the United States; that the educational system yields individuals who can meet the needs of the Nation's science, mathematics, and engineering work force; and that the public's awareness of, and interest in, scientific and technological developments is strongly supported.

Currently, dozens of EHR-administered programs, as well as a number of science and mathematics education projects that the Foundation shares with other government agencies, are being assessed; full evaluations for the entire EHR portfolio are to be completed within the next several years.

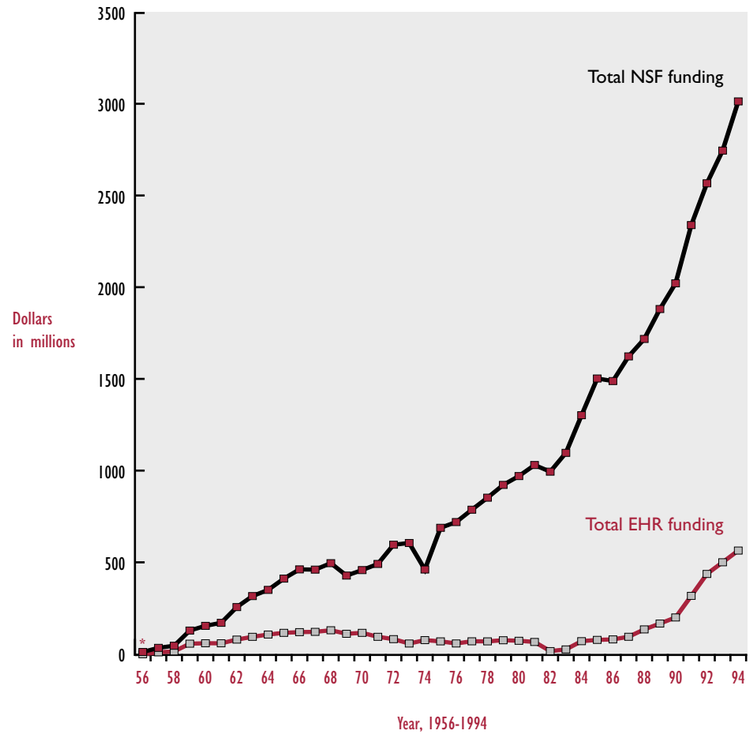
The evaluation initiative was launched by EHR in response to a growing awareness—building through the 1980s and eventually emerging as a mandate for action—that the attention being paid to monitoring EHR's science and mathematics education programs was deficient. The Directorate's approach to monitoring its programs needed to be revised. Traditionally, programs had been judged largely according to their impact on individual grantees—reports, for example, of an awardee's satisfaction with a particular program or records of career advances or scientific achievements that could be attributed to his or her NSF affiliation.

By 1992 it was clear that a broader based, outcomes-oriented process of evaluation had to be implemented. As EHR's program portfolio expanded, administrators had to have richer, fuller data to monitor the progress of their programs, to strive for their continuous improvement, and to measure their gains over time. No longer was it sufficient to know how an individual grant awardee had used his or her funding and what specific benefits had resulted. Rather, it was decided, evaluations must reveal, among other things, the overall progress of a program in relation to its stated mission.

EHR today is required to account to Congress and the Administration for the effectiveness of its programmatic activities.

Also by 1994, EHR's annual federal financial support had grown dramatically—to $569 million from $71 million in 1981 (see Exhibit 1). As the funding mounted, Congress became increasingly intent on keeping a closer eye on where, exactly, tax money was going and how effectively EHR was using it. With passage of the Government Performance and Results Act of 1993, accountability based on evaluation of a program's consistent results was no longer merely desirable—it was the law. Simply stated, EHR today is required to account to Congress and the Administration for the effectiveness of its programmatic activities.

Exhibit 1: Funding for NSF and for the Directorate for Education and Human Resources (EHR), 1956-94



* 1956 Total NSF funding = 16 million
1956 EHR funding = 3.5 million

Source: Indicators of Science and Mathematics Education 1995

Since 1992, developing and implementing thorough and insightful EHR program evaluations has been the responsibility of the Directorate's Division of Research, Evaluation and Communication (REC), which was formed in 1991.

To fulfill its evaluation mandate, REC has for the past 4 years pursued three broad objectives:

- To design and oversee periodic evaluations of the approximately 30 programs currently in EHR's portfolio;
- To elevate the practice of science and mathematics education evaluation by developing new and better methods; and
- To build, through workshops and other means, the capacity of NSF, other government agencies whose missions include science education, and individual grant awardees to conduct rigorous evaluations.

In addition to implementing evaluations of EHR programs, REC has during the past several years engaged in evaluation activities associated with science, mathematics, engineering, and technology education programs administered by other federal agencies.

To carry out its evaluation work, the REC staff focuses on developing three kinds of studies:

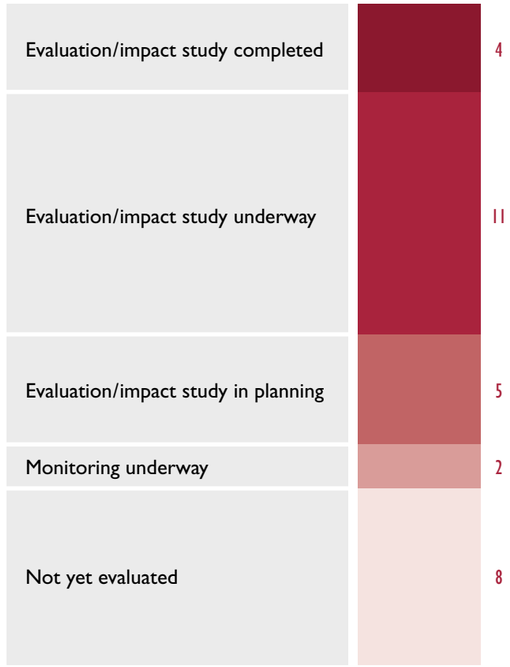
- Evaluations: systematic examinations conducted by external evaluators to determine the merit or worth of programs and ways in which they can be improved;

- **Impact studies:** briefer, quicker examinations of programmatic effectiveness that yield reports more limited in their focus, data collection, and analysis than a full evaluation; and

- **Monitoring:** ongoing collection and analysis of data on the status of EHR projects.

A staff of REC evaluation officers oversees the design and execution of evaluations by independent contractors, who are retained to ensure the objectivity with which a program is examined. Most of these studies take up to 2 years to complete, at a cost ranging from $100,000 to $1 million. In fiscal year 1994, program evaluation represented an EHR investment of about $12 million—roughly 2 percent of the Directorate's total budget. The Directorate has placed all 30 of its programs on a 5-year evaluation cycle, so that at least one-third of the portfolio is being evaluated at all times (see Exhibit 2).

Exhibit 2: Evaluation Status of EHR's 30 Programs



At the conclusion of the evaluation process for a given program, REC submits the contractor's final report to the NSF Assistant Director in charge of EHR, along with a set of preliminary recommendations suggesting actions with respect to the program based on the findings concerning its effectiveness. The main purpose of the evaluation—apart from its function as an accountability document—is to help the Assistant Director make management decisions regarding the Directorate's program portfolio, identify ways in which programs might be redirected or consolidated, and detect emerging science, mathematics, engineering, and technology priority areas that may merit more programmatic attention.

The Directorate has placed all 30 of its programs on a 5-year evaluation cycle.

Since 1992, evaluations of varying scope have been completed for four programs; more than a dozen others are either in progress or in the planning stage. Additionally, REC has played a leadership role in several other evaluation-related projects, including interagency program evaluations and how-to documents focusing on the evaluation process itself.

THE VALUE OF EVALUATION

HIGHLIGHTS

EVALUATION-RELATED ACTIVITIES

FUTURE EVALUATION EFFORTS

HIGHLIGHTS

Recognizing the value of program evaluations as internal management tools, and responding to Congress's demand for them as accountability documents, EHR is moving forward with its evaluation effort.

The reports published to date, examining the efficacy of nine EHR programs—differ from one another in purpose, scope, level of detail, and length. Several evaluations, for example, have yielded extensive reports, intended to serve as cumulative assessments of their subjects' progress over an extended period of time; others are comparatively brief "impact" studies, intended to inform NSF administrators of their subjects' progress; and one is a preliminary report of a full-scale evaluation that is scheduled for completion in 1996. Following are highlights of the reports published thus far—descriptions of the programs that have undergone evaluation, accounts of the evaluation process associated with each, and a brief sampling of findings arising from the evaluation.

Statewide Systemic Initiatives Program (SSI)

Description. SSI was launched by EHR in 1991 as a major endeavor toward improving science and mathematics education for all children in the United States through comprehensive, state-level systemic reform. The goal of the program is to move from independently devised reform efforts to state-initiated measures involving the coordinated improvement of many aspects of the education system, including teacher preparation, instructional materials, and assessment of student learning.

SSI states take seriously the idea that a vision of good practice must guide their efforts.

To seed and scale interventions in the teaching and learning of science and mathematics, EHR participates in funding arrangements called "cooperative agreements"—awards made to states of up

to $2 million per year for 1 to 5 years for the purpose of forging partnerships among educational institutions that will lead to ambitious, coherent, and comprehensive approaches to statewide reform. To be eligible, states must demonstrate both commitment to reform and the support of its government officials and educators, since such support is an essential element in the success of science and mathematics education reform.

In 1991, an initial competition resulted in cooperative agreements with 10 states; a second competition in 1992 resulted in 11 additional awards; and a third competition in 1993 resulted in awards to five more states (see Exhibit 3).

Evaluation. In 1992, NSF contracted with a consortium of research organizations led by SRI International of Menlo Park, California, to conduct a 5-year evaluation to examine individual projects and determine how, in the aggregate, they are performing to promote and sustain systemic reform. During the first year, the evaluation team concentrated on descriptive activities, collecting data on each funded state's approach to systemic reform and its plan for implementation. During the second year, the evaluators focused on the funded states' accomplishments in three areas: good classroom practice; equity (access for all students to high-quality science and mathematics education); and coordination of reform throughout their respective systems.

The methodology in years 1 and 2 of the evaluation combined an extensive review of documents (including annual reports from the funded states to NSF) with supplemental telephone interviews with state officials. Also, indepth case studies — including week-long visits — were conducted in nine states selected to represent a variety of educational and demographic conditions and varying approaches to systemic reform. In addition, site visits were conducted in three non-case-study states. To date, the following SSI evaluation reports have been published:

- *Evaluation of the National Science Foundation's Statewide Systemic Initiatives (SSI) Program: First Year Report, Volume I, Technical Report (SRI International 1994).*

- *Evaluation of the National Science Foundation's Statewide Systemic Initiatives (SSI) Program: Second Year Report: Cross-Cutting Themes (SRI International 1995).*

- *Evaluation of the National Science Foundation's Statewide Systemic Initiatives (SSI) Program: Second-Year Case Studies: Connecticut, Delaware, and Montana (SRI International 1995).*

Findings. Given the wide latitude that states have in implementing systemic reform, generalizations about program emphases and activities vary considerably, but some common elements can be observed. For example:

- SSI states take seriously the idea that a vision of good practice must guide their efforts. Many are developing new statewide curriculum frameworks, and there is growing agreement that a central goal of mathematics and science education is to promote critical thinking, conceptual understanding, and problem-solving ability.

Exhibit 3: Recipients of Statewide Systemic Initiatives (SSI) Awards, 1991-93

1991 Awardees	1992 Awardees	1993 Awardees
Connecticut	California	Arkansas
Delaware	Georgia	Colorado
Florida **	Kentucky	New Jersey
Louisiana	Maine	New York
Montana	Massachusetts	South Carolina
North Carolina	Michigan	
Nebraska	New Mexico	
Ohio	Puerto Rico	
Rhode Island *	Texas	
South Dakota	Vermont	
	Virginia **	

* cooperative agreement discontinued 1994
** cooperative agreement discontinued 1996

- All SSI states are seeking to develop and articulate clear goals for what students should know and be able to do. These goals are fairly uniform across states for mathematics, but less so for science.

- Reform strategies vary greatly by state, but a shared emphasis has been on inservice training of teachers (see Exhibit 4). More than one-third of SSI funds awarded are allocated to this activity, and nearly 50,000 teachers received training (see Exhibit 5).

- All SSI states have developed strategies for improving educational equity by providing access for all students to high-quality science and mathematics education. The strategies and the groups targeted for service vary; most commonly, the focus is on women and minorities.

- Sources other than SSI are providing (in the aggregate) more than a dollar-for-dollar match to NSF's investment.

Exhibit 4: Statewide Systemic Initiatives (SSI) Change Strategies: Findings and Problems

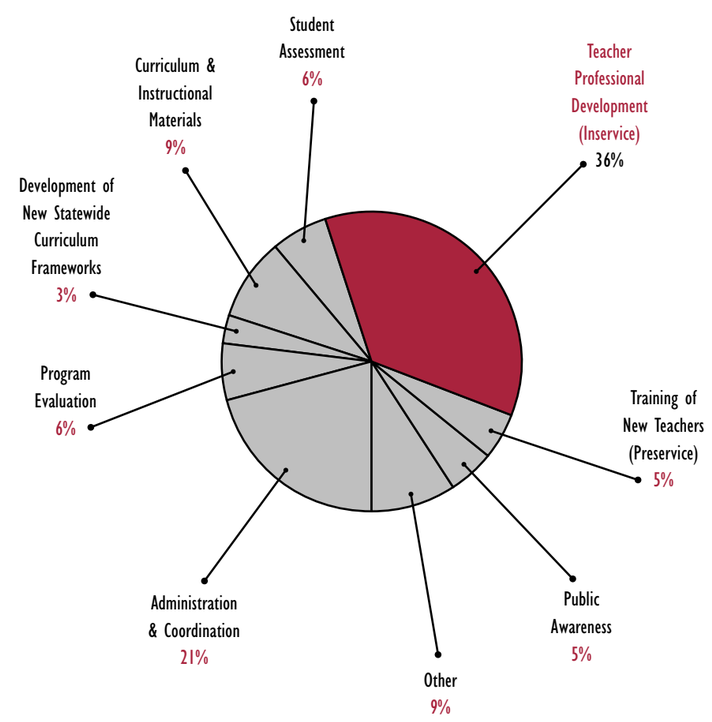
Findings

- SSIs spent the greatest portion of their NSF dollars on the professional development of practicing mathematics and science teachers, providing such services to nearly 50,000 teachers during the past year (about 8% of the public school teachers in SSI states).
- NSF's funds leveraged $83 million from other sources for the SSI program.
- SSIs use a wide combination and variety of reform strategies, ranging from aligning state policies to supporting model schools to developing new instructional materials.
- Many SSIs are working to strengthen the education infrastructure to support the long-term improvement of mathematics and science instruction beyond the end of the SSI program.
- Three states target the introduction of new instructional materials as a primary SSI strategy; many others disseminate information about existing materials that they consider effective.
- Five states support model schools in their reform strategy.
- Ten states target local districts and/or communities as part of their SSI change strategy.

Problems and Issues

- A number of state policies, such as those for assessment of student learning, often remain out of alignment with states' reform goals.
- The SSIs give little attention to local education policy systems.
- A strength of the SSIs is their focus on building the capacity of practicing teachers to improve the teaching of mathematics and science, yet the level of effort is often insufficient to meet the challenges of reform.
- States are paying relatively little attention to preparing the next generation of teachers.
- SSI states most often rely on persuasion and incentives to motivate educators to participate in their reform agenda, although there also have been real changes in authority relations in a few states.
- The task of mobilizing public opinion is a difficult one and often requires skills not readily available within the mathematics, science, and education communities promoting the reform agenda.

Exhibit 5: Statewide Systemic Initiatives (SSI) Expenditure of NSF Funds in FY 1993 by Purpose of Activity



On the other hand, progress in achieving some goals may be lagging, and states may have to revise some of their strategies. For example:

- NSF has targeted equity as a key objective, but while the participating states have adopted various strategies to accomplish this goal, some have not yet developed coherent plans.

- States are finding the need to mobilize public support for systemic reform greater than was anticipated and, therefore, must increase efforts to attract support.

- The alignment of all state policies to achieve systemic reform is proving difficult. In particular, technical, financial, and political barriers to changing methods of assessing student learning present a major impediment to change.

- Issues that have largely been neglected to date include state strategies for transferring lessons from model sites to schools; attention to local education systems, including school districts and boards; and attention to the need for changes in the preparation of the next generation of teachers.

Overall, while measuring systemic change and assessing the impact of a broad program such as SSI (with its large expenditures and multiplicity of jurisdictions) is a complex undertaking, this evaluation has played an important role in providing NSF and its partners with a useful indication of education reforms introduced in classrooms, schools, districts, and across states.

The SSI evaluation has played an important role in providing NSF and its partners with a useful indication of education reforms.

NSF is implementing a number of changes in response to the year 1 and 2 evaluation reports, including the provision of a full range of technical assistance services. To this end, the technical assistance provider is helping states develop strategies to inform the public about what has been accomplished by these reform efforts and to share lessons with other NSF systemic initiatives such as those in urban and rural areas. In addition, the evaluation has redoubled efforts to expand the technology component of teacher professional development, to stimulate NSF strategic planning in teacher preparation, and to sustain reform activities in states where the SSI awards will soon expire.

Presidential Awards for Excellence in Science and Mathematics Teaching Program (PAESMT)

Description. PAESMT was established in 1983 under the sponsorship of NSF's Division of Elementary, Secondary and Informal Education with the goal of providing national recognition to outstanding elementary and secondary school teachers in science and mathematics. The program also seeks to encourage active leadership roles in science and mathematics education by former awardees.

Each year, after a lengthy nomination and selection process, four mathematics and science teachers—two elementary and two secondary—are chosen for the award in every state and U.S. territory. Each award is accompanied by an NSF grant of $7,500, to be used at the teacher's discretion to improve science or mathematics education at the awardee's home institution. The awardees are brought to Washington, D.C., for a week of activities and professional interaction.

PAESMT program awardees reported a renewed sense of validation for their efforts and reinforcement of their motivation to continue to teach.

During the first decade of the program's existence, more than 1,600 teachers from around the Nation received national recognition through this program for their subject-matter expertise and teaching skills.

Evaluation. From the time of its inception through 1994, no evaluation of this program had been undertaken. However, in 1994, the Foundation contracted with Westat, Inc., to engage in a small-scale study focusing primarily on the personal and professional impacts of the award. Telephone conversations were conducted with 115 of the 430 teachers who had received awards in 1990 and 1991. An interview protocol was developed to guide the conversations, but the exact questions varied among respondents, and the awardees were given considerable latitude in shaping the content of the interview.

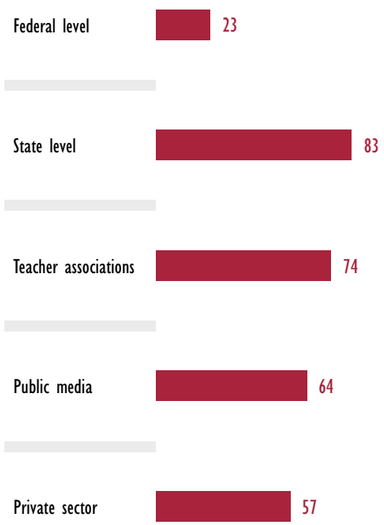
In addition to the impact of the award on recipients, other topics that arose during the phone interviews included their use of the NSF grant money and additional financial support that many of the awardees received from other donors in connection with the NSF award. The evaluation also sought awardees' views of the nomination and selection process and their suggestions for strengthening the program and broadening its coverage.

This study resulted in a publication covering all aspects of the evaluation:

> *Short-Term Impact Study of the Presidential Awards for Excellence in Science and Mathematics Teaching (Westat 1994).*

Findings. PAESMT was found to be an overwhelming success in terms of its impact on participants and the recognition it provides to the importance of good mathematics and science instruction. Strong positive effects were found at the professional and personal levels for those who have been recognized, with awardees reporting a renewed sense of validation for their efforts and reinforcement of their motivation to continue to teach. Within an awardee's local community, there was much pride and greater interest in effective mathematics and science instruction on the part of students, parents, and other teachers. A majority of awardees reported that the award increased their opportunities to make improvements in their schools' mathematics and science programs; many of them became involved, following their recognition by NSF, in curriculum development and the supervision of teacher-training workshops. Opportunities for involvement, even leadership, at the state level—through participation on panels and committees dealing with education—also had increased for some grant recipients (see Exhibit 6). In a number of cases, the award and the publicity associated with it led to gifts or awards from other organizations.

Exhibit 6: Percent of Presidential Awards for Excellence in Science and Mathematics Teaching (PAESMT) Program Awardees from 1990 to 1991 Reporting Positive Effects on Opportunities for Leadership at Various Levels, 1994



The financial award ($7,500) was seen as an important part of the program by 90 percent of the interviewees. It was used in a wide variety of ways, with most recipients spreading the funds among a number of innovative and appropriate uses—sharing the reward, for example, with colleagues for their projects, or providing support for conference attendance and training activities.

PAESMT was found to be an overwhelming success in terms of its impact on participants and the recognition it provides to the importance of good mathematics and science instruction.

Negative comments were received by the evaluators regarding only two aspects of this program. The first had to do with the application process, described by several interviewees as so burdensome that it might discourage some worthy candidates. A number of interviewees also felt that the existence of the program and the nomination process are only known to persons active in professional associations. They recommended increased dissemination of information concerning the program to bring about wider participation by all segments of the teaching profession. In response to these comments, NSF is considering ways in which the application process can be simplified and awareness of the program increased.

Overall, the evaluators summarized their assessment of PAESMT by calling it a program with significant impact on the teachers recognized; these teachers, in turn, have impacted their colleagues in some very significant ways.

Young Scholars Program (YSP)

Description. Launched in 1988 by NSF's Division of Elementary, Secondary and Informal Education, YSP typically funds summer projects of from 3 to 8 weeks duration in universities and other organizations that conduct scientific research. Junior and senior high school students are eligible to participate, the goal of the program being to excite students about science, mathematics, and technology and to encourage them to investigate and pursue careers in these fields. Related

Preliminary findings suggest that the YSP has been especially encouraging or reinforcing for African American students and least encouraging or reinforcing for female participants.

goals are to acquaint them with various career options in the sciences, engineering, and technology; to increase their awareness of the academic preparation needed for such careers; to acquaint them with academic and research environments; and to contribute to their confidence in their ability to make career decisions. The program strongly emphasizes student participation in the process of scientific discovery through interaction with practicing scientists in the laboratory and other research environments.

In recent years, the program has supported more than 140 projects each year, involving more than 6,000 students annually at sites throughout the United States. Supporting NSF grants are awarded to colleges, universities, or other organizations and institutions whose members are primarily university faculty members or researchers.

Evaluation. The first formal evaluation of this program, a short-term impact study, was conducted by Westat, Inc., in 1994. The study was intended to be small-scale and exploratory, designed primarily to obtain information about participating students' perceptions of their experiences with the program, about their educational and career plans, and about the perceived impact that their YSP experiences had on those plans. In addition to student self-reports, evaluators sought the opinions of parents. Also, information about interest in science and career choices was obtained from students who applied to, but did not attend, a YSP project.

The study indicates that the YSP experience was a very positive one for the overwhelming majority of participants.

The evaluators conducted telephone interviews with 215 participants and 70 nonparticipants, all of whom had participated or applied in 1991. Interviews were also conducted with 52 parents. One publication was issued in connection with this evaluation:

Short-Term Impact Study of the National Science Foundation's Young Scholars Program (Westat 1994).

Findings. The study indicates that the YSP experience was a very positive one for the overwhelming majority of participants. They gained awareness of a science-related community of professionals, learned a good deal about various fields, and discovered much about themselves and their professional interests. Of course, given the program's selection process, which seeks to identify students of high ability who have an interest in science, mathematics, or engineering, it follows that the great majority of participants entered the program predisposed and qualified to benefit from it.

Among the high school seniors in the 1991 program—almost all of whom were still in college at the time of the 1994 evaluation—65 percent were majoring in science, mathematics, or engineering, and 24 percent were majoring in health-related fields. However, data also revealed similar college choices among students who had applied but did not participate in the 1991 program.

Findings suggest that the program has been especially encouraging or reinforcing for African American students (see Exhibit 7). It was

Exhibit 7: Percent of 1991 Young Scholars Program (YSP) Participants Reporting Various Career Plans, by Race, 1994

Exhibit 8: Percent of 1991 Young Scholars Program (YSP) Participants Reporting Various Career Plans, by Gender, 1994

least encouraging or reinforcing for women in their decisions to pursue careers in the sciences and engineering (see Exhibit 8, previous page). Eighty-one percent of African American participants reported career intentions in science, mathematics, engineering, and technical (SMET) fields. Just 33 percent of female participants indicated such a choice.

The study raised a number of issues concerning the program that should be considered in the future. Chief among these is the advisability of structuring the program to ensure its attractiveness and availability to young women, members of racial minorities, and students of high potential as well as those whose high ability has been established. In response, NSF evaluation staff has recommended a followup study to provide more information about how the program can provide greater, more positive impacts on these students, especially the young women.

Informal Science Education Program (ISE)

Description. Since the early 1980s, the ISE has functioned with the goal of advancing science learning for the general public. The program, administered by EHR's Division of Elementary, Secondary and Informal Education, is devoted to supporting public and commercial television and radio programming, the creation of films and videos, museum exhibits, science and technology center exhibits, professional development activities, and community- and youth-oriented organizations.

The program was designed to provide rich and stimulating environments—outside of the school setting—where individuals of all ages, backgrounds, and interests can increase their appreciation and understanding of science, mathematics, and technology. ISE's goals have been sharpened in recent years with an emphasis on establishing linkages between formal and informal education and increasing the number of young people involved in science-related activities, especially those from minority or otherwise underserved groups.

Nearly 80 percent of ISE funds went to science and natural history museums, zoos, and children's museums and to media organizations, such as production companies and television stations.

Evaluation. In 1994, NSF contracted with Cosmos of Bethesda, Maryland, to conduct an evaluation of the ISE program, which is to be completed in 1996. Designing the evaluation for a program with activities and goals that are highly diverse and do not lend themselves to direct impact measurement of outcomes associated with a given activity is a complex undertaking.

The evaluation methodology developed by the contractor combines surveys, case studies, and 21 site visits (15 to ISE-funded projects sites and 6 to science education project sites not funded by the program). Interviews with past and present project directors and other knowledgeable sources also are planned. In addition, a meta-analysis of existing pertinent evaluations is planned. As a first step, the evaluator inventoried and analyzed the types of activities and institutions supported by ISE. In addition to technical reports dealing with the evaluation plan, a report summarizing some preliminary findings has been issued:

ISE-funded projects were especially influential in providing access to science education for previously underserved populations and in introducing fields of science to the public.

Informal Science Education Program: Evaluation Design Brief and Report of Preliminary Findings (Cosmos 1995).

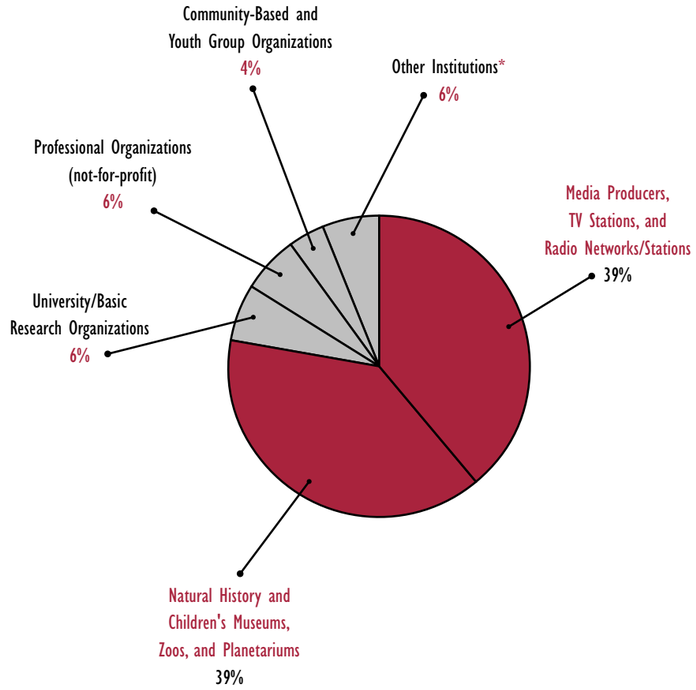
Findings. From 1984 to 1995, funding for ISE increased from less than $5 million to $35 million. Of the total grant funds awarded ($183 million), more than three-fourths went to support museum exhibits and the production of television programs. The remainder of the funds was awarded to a variety of activities, including after-school and community programs, radio, and film production. Nearly 80 percent of the funds went to science and natural history museums, zoos, and children's museums and to media organizations, such as production companies and television stations (see Exhibit 9, page 26).

Preliminary survey data show that young people involved in informal science activities were predominantly between the ages of 12 and 17 and that NSF-funded projects were especially influential in providing access to science education for previously underserved populations and in introducing fields of science to the public. The case studies offer evidence that program goals are being met: For example, in one California county, two eighth grade girls have

taken the initiative to develop hands-on geology materials after participating in an ISE activity run by the University of California at Davis, a result that relates to the ISE goal of promoting science interest in students from groups that are traditionally underrepresented in science—in this case, girls.

Another stated ISE goal—that of establishing new relationships between formal and informal science education—was addressed in Michigan, where a school district conducted a summer program using "The Magic School Bus," an ISE-funded project that produced a popular book and television series. Furthermore, videos dealing with environmental issues that were produced with ISE support

Exhibit 9: Types of Institutions Receiving Support from Informal Science Education Program, 1984-95 (Total = $183 million)



*Other institutions include aquariums; botanical gardens and arboretums; for-profit organizations; general museums; government agencies; and nature centers.

were seen by two other federal agencies—the Department of the Army and the Substance Abuse and Mental Health Services Administration—which, in turn, commissioned additional videos and activity books for distribution in schools. In that case, the ISE program goal of establishing linkages among a variety of organizations was addressed.

The evaluators' final report will present the first summative analysis of the NSF investment in ISE.

Instrumentation and Laboratory Improvement Program (ILI)

Description. ILI was launched in 1985 by NSF's Division of Undergraduate Education (DUE), whose mission is to strengthen and ensure the vitality of undergraduate education in science, mathematics, engineering, and technology. The ILI program (one of three DUE efforts on which reports have been completed) supports projects aimed toward generating new and improved approaches to laboratory- and field-based instruction. Above all, the program is designed to stimulate student interest in science- and technology-related courses and degree programs and to support the development of national models for the improvement of undergraduate laboratory instruction.

ILI has become one of the Foundation's largest and most visible programs in the area of science education, attracting more than 2,000 proposals annually and awarding about $20 million a year. From its inception in 1985, it has made more than 4,700 grants to 1,200 different educational institutions. And since grantee institutions are required to match ILI funds by at least 100 percent, the program has generated an additional $316 million in support of laboratory improvement.

> Matching funds were frequently generated by private and industrial sources that regard an institution's winning an NSF ILI grant as a sign of quality assurance.

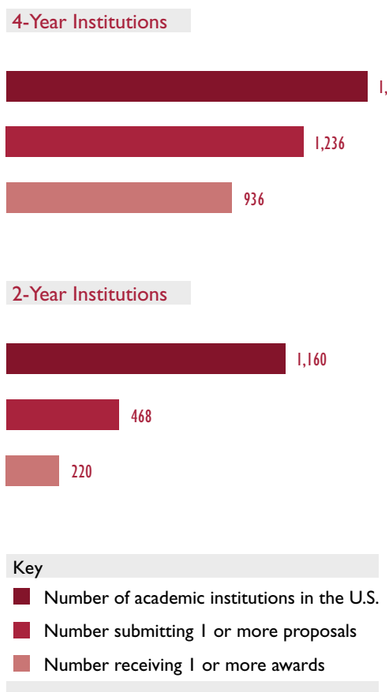
Evaluation. In October 1994, the Foundation began conducting a full-scale evaluation of ILI, scheduled for completion by Westat, Inc., in 1996. It will assess how successful the program has been in reforming undergraduate laboratory instruction and in meeting the needs of various kinds of institutions and individual disciplines. The evaluation also seeks to address the effect of the program on curricula, faculty

teaching practices, and on the training of future elementary and secondary teachers.

The evaluation covers the period 1985-94. Data sources include the analysis of existing ILI records; a mail survey of ILI grantees from 1990 and 1992; a survey of individuals whose ILI applications were unsuccessful; "tracer studies" documenting the impacts of innovative projects developed by grantees on other institutions; and site visits to selected colleges and universities. To date, the ongoing ILI evaluation has yielded a preliminary document:

- *A Short-Term Impact Study of the National Science Foundation's Instrumentation and Laboratory Improvement Program (ILI) (Westat 1996).*

Exhibit 10: Number of Postsecondary Institutions That Applied for and Received Instrumentation and Laboratory Improvement (ILI) Program Awards, 1985-94



Findings. The preliminary report answers one important evaluation question: Does the program reach the intended audience? The program has been generally successful in this respect. Most important, undergraduates from both 4-year and 2-year institutions who major in NSF-supported disciplines are being reached by the program; and 89 percent of bachelor's degree recipients in these fields studied at institutions that received one or more ILI awards (see Exhibit 10).

Information gathered during early site visits points to other noteworthy program outcomes. For example, matching funds were frequently generated by private and industrial sources that regard an institution's winning an NSF grant as a sign of quality assurance; many students have pursued advanced science/technology studies or careers as a result of their experiences in ILI-enabled laboratories; and the presence of ILI-funded laboratory equipment has attracted new faculty at some institutions.

Criticisms expressed to evaluators about the program concerned such matters as the need for improved communication regarding ILI (preferably through the use of electronic media such as the Internet); the frustration experienced by principal investigators owing to time constraints during the early stages of installing ILI-funded equipment; and the lack of funding for maintenance and repair of equipment purchased under an ILI grant.

Many students have pursued advanced science/technology studies or careers as a result of their experiences in ILI-enabled laboratories.

Course and Curriculum Development Program (CCD)

Description. CCD, another program funded by DUE, awards grants to universities for developing undergraduate science and mathematics units, courses, and sequences of courses. Its aim is to support major instructional changes at the undergraduate level that have potential national impact. Two major goals are pursued:

- To increase the understanding, interest, and comfort of students engaged in science- and mathematics-related studies, especially those who are members of minorities and other groups traditionally underrepresented in those fields; and

- To contribute to a shift in academic culture that will foster the placement of greater value on undergraduate teaching and learning.

This program especially encourages development of innovative introductory-level science and mathematics courses for both majors and nonmajors. Grants provide for the implementation, assessment, and dissemination of projects designed to improve curricula and the learning environment and to develop new materials, software, and technologies, as well as courses.

Substantial changes have occurred in the thinking and behavior of the great majority of faculty members in CCD participating departments.

The first CCD grants were awarded in 1988. For the first 3 years, awards were made for projects focusing on calculus, precalculus, and engineering.

Evaluation. An evaluation of this program, covering its first 5 years (1988-93), is currently being conducted by Network, Inc., of Andover, Massachusetts. During this 5-year period, approximately 400 awards were made. The evaluation methodology consists of three major components:

- Surveys of all principal investigators who received awards and a sample of individuals who submitted proposals but were not funded;
- Site visits to 27 grant-recipient institutions; and
- Twelve case studies.

A final report is due in the summer of 1996. An interim report was issued in October 1995, based on partial data available at that time:

- *Evaluation of the Division of Undergraduate Education's Course & Curriculum Development Program, Interim Report (The Network, Inc. 1995).*

Findings. The evaluators have specified four questions that they are seeking to answer with data from the surveys and site visits. Two of the questions deal with project impact on faculty, students, and departments; and two questions deal with barriers that interfere with success and the ways in which successful projects have overcome these barriers. The interim report focused chiefly on the first two questions.

The study revealed that substantial changes have occurred in the thinking and behavior of the great majority of faculty members in participating departments. Project directors generally felt that compared to their preaward instructional methods, faculty members were relying less on lecturing, were using new instruction materials and new methods of student assessment, and were spending more time teaching. At the same time, students were more often working in teams, developing their own research questions, and using instruction-related software.

Barriers to CCD program success include the traditional placement of higher value on "substantive" research activities than on research into teaching practices.

Project directors also saw marked gains in students' understanding, competence, and comfort in dealing with science and mathematics subjects (see Exhibit 11). The site visits yielded evidence that the greatest gains were often experienced by students who had difficulty in the past with traditional instruction practices. Although, according to the interim report, the CCD program was not yet having a profound impact on departmental policies (such as increased commitment to undergraduate education), the program was having significant success in vitalizing undergraduate teaching and encouraging new approaches to learning.

Regarding barriers to student gains, some students had difficulty adopting new learning behaviors and skills, which some projects attempted to remedy through appropriate course materials or tutorial centers. Barriers to project implementation included psychological, cultural, and structural factors: for example, the evaluators found that many faculty members enjoy their roles as "sources of wisdom" in traditional teacher settings and, thus, are reluctant to play what they may view as a mere "facilitator-of-learning" role. Other barriers to program success include the traditional placement of higher value on "substantive" research activities than on research into teaching practices and the reluctance of some university departments to allocate funds for nontraditional programs.

The evaluators suggested a number of strategies for overcoming these barriers to implementation. They plan, in their final CCD evaluation report, to recommend ways in which the program will continue to be successful in the future.

Exhibit 11: Four Areas in Which 1988-93 Project Directors Felt That Students Benefit from Projects, 1995

	Percent of project directors holding this opinion
Student gained competence in applying concepts, principles, or theories	81
Student gained competence in using methods and equipment	76
Student gained understanding or familiarity with the scientific approach to problems	72
Student developed greater interest in or comfort with the science taught	71

Undergraduate Faculty Enhancement Program (UFEP)

Description. This program was inaugurated by EHR's DUE in 1988 in response to a 1986 National Science Board task force report addressing undergraduate education in mathematics, engineering, and the sciences. The report recommended that NSF establish "a comprehensive set of programs to catalyze and stimulate national efforts to assure a vital faculty" in these areas of study. Accordingly, UFEP was designed to assist faculty members who are primarily engaged in the instruction of undergraduates to gain experience with recent advances in their fields, with new experimental techniques, and with ways of incorporating these into undergraduate instruction. Projects are regional or national in scope and typically consist of hands-on short courses or workshops, along with followup activities that encourage sustained interaction among participants. A major component of the program is regional coalitions between 2- and 4-year institutions.

Both participants and project directors expressed enthusiasm for the UFEP projects in which they were involved, finding them either highly valuable or, at least, very worthwhile.

Proposals are accepted from any organization with the scientific expertise and facilities to conduct these projects. Subject matter of faculty enhancement activity may pertain to any field of science, engineering, or mathematics normally supported by NSF. During its first 3 years of operation (1988-90), the program awarded approximately $6.7 million to 92 projects, each of which supported one or more workshops or short courses and served nearly 3,000 participants.

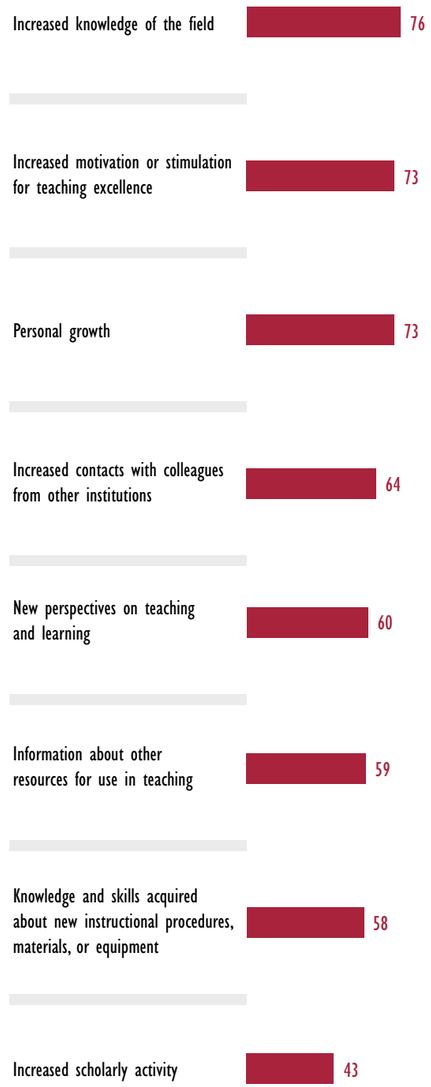
Evaluation. To assess the progress of UFEP, NSF contracted with Westat, Inc., of Rockville, Maryland, to evaluate the program's performance during its first 3 years. The evaluation was designed to consider effectiveness mainly in three areas: (1) Did the projects that were funded meet the participants' needs? (2) Did the program as implemented meet the needs of the profession? (3) Were the programs and criteria as defined by NSF appropriate to meet program goals?

In approaching the evaluation, Westat obtained data through questionnaires completed by directors of 91 of the 92 projects that had been funded between 1988 and 1990, and from telephone interviews with a sample of 469 participants. Additional information for the evaluation was obtained by an NSF-appointed Assessment Advisory Committee by means of discussions with groups of faculty members and participants in the course of professional society meetings. Three publications reporting on the UFEP evaluation were issued by NSF in 1993:

- *Assessment of the National Science Foundation's 1988-1990 Undergraduate Faculty Enhancement Program: Final Report (Westat 1993).*
- *Assessment of the National Science Foundation's 1988-90 Undergraduate Faculty Enhancement Program: Executive Summary (Westat 1993).*
- *Assessment of the National Science Foundation's Undergraduate Faculty Enhancement Program: Interpretive Overview. A Statement from the Assessment Advisory Committee (Westat 1993).*

Findings. From the vantage point of both participants and project directors, the first objective—meeting participants' needs—was effectively met. Both groups expressed enthusiasm for the projects in which they were involved, finding them either highly valuable or, at least, very worthwhile. As benefits derived from the projects, respondents singled out increased knowledge of their field, personal growth or renewal, increased motivation to aspire for excellence as teachers, and increased contacts with colleagues from other institutions (see Exhibit 12).

Exhibit 12: Percent of Faculty Enhancement Program (UFEP) Participants from 1988 to 1990 Who Selected Above-Average Ratings When Asked About Specific Benefits Derived from Their Experience with the Program, 1993



The majority of participants reported modifying their teaching methods, introducing new content, and acquiring new equipment.

Findings were less clear-cut regarding two other questions that the evaluation sought to answer: Is the program furthering the needs of the science and mathematics teaching professions in general? And is the program structured suitably to meet its goals? When asked to assess the impact of the program on their profession, educational organization members and people holding faculty positions varied in their responses, depending on their disciplines. Chemistry instructors, for example, were very affirmative, while physics instructors were considerably less so. And concerning UFEP goals, the Assessment Advisory Committee expressed some concerns: participants included a high proportion of individuals who, even without program encouragement, were already very active professionally, judging from their abundant publications as well as their frequent attendance at meetings, seminars, and workshops. Furthermore, the committee concluded that UFEP administrators might be able to do a better job of encouraging the participation of women and minorities in the program.

The majority of UFEP participants reported modifying their teaching methods, introducing new content, and acquiring new equipment.

Overall, however, the evaluation indicated that the program, despite its weaknesses, has been of some value in faculty development and renewal. The Assessment Advisory Committee recommended that a wider net be cast in recruiting participants, especially women, minorities, and others who are not among the most professionally active members of their profession.

In response to the evaluation, NSF made several changes to the UFEP funding guidelines. First, to ensure the participation of those faculty who are most in need of professional development, NSF now strongly encourages UFEP proposers to solicit the participation of faculty at 2-year institutions and faculty just starting their academic careers. Second, the guidelines were modified to place more emphasis on the participation of women and members of underrepresented racial and ethnic groups in UFEP workshops. Finally, the eligibility requirements of the program were expanded to include scientific societies and associations as UFEP host organizations.

Visiting Professorships for Women Program (VPW)

Description. The VPW program was initiated by EHR's Division of Human Resource Development in 1982 to address the underrepresentation of women in science. It supports experienced women scientists and engineers who serve as visiting faculty members at host institutions, enabling these visiting professors to do research at an academic institution of their choice, while also serving as role models for younger women who may aspire to careers in science, mathematics, engineering, or technology. In addition to research and teaching, the visiting professor must undertake counseling of students and other interactive activities to increase the visibility of women scientists in the academic environment and to provide encouragement to other women to pursue careers in science and engineering.

Approximately 90 percent of awardees rated VPW as having a positive impact on their professional development, research activity, research career generally, and scientific reputation.

Awards are usually granted for periods of 6 to 15 months, with recipients expected to spend approximately 70 percent of the period on research activities and 30 percent on teaching and other interactive activities. Between 1982 and 1992, approximately 100 women applied each year, and 29 percent of them received awards.

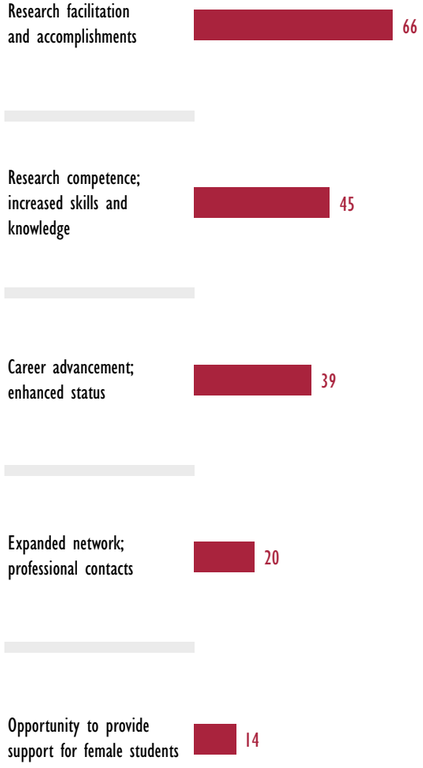
Evaluation. The Foundation contracted with SRI International at the close of the program's first decade to evaluate the outcomes of the visiting professorship program and also to identify the barriers that women perceive as retarding their progress in science and engineering careers. Thus, the evaluation was a two-phase undertaking—the first half dealing with career advancement for women in science, the second focusing on visiting professorship experiences of the awardees and evidence of program effects.

In conducting the evaluation of the program's effectiveness, SRI International gathered quantitative and qualitative data through mail questionnaires. Respondents included more than 200 VPW awardees and, for purposes of comparison, more than 300 applicants who did not receive an award (declinees), plus more than 300 women who were recipients of NSF grants other than those offered by the visiting

professorship program. The study was completed in 1993; two publications have resulted from it:

- *The Visiting Professorships for Women Program: Lowering the Hurdles for Women in Science and Engineering: Final Report (SRI International 1993).*
- *The Visiting Professorships for Women Program: Lowering the Hurdles for Women in Science and Engineering: NSF Summary and Comments (SRI International 1994).*

Exhibit 13: Percent of Visiting Professorships for Women (VPW) Program Awardees from 1982 through 1992 Reporting Various Types of Program Benefits They Experienced, 1993



Findings. The study was primarily devoted to assessing the first program goal: providing advancement opportunities for participants. The evaluation, based on participants' self-reports, documents a number of benefits. Approximately 90 percent of awardees rated VPW as having a positive impact in four areas: their professional development, research activity, research career generally, and scientific reputation. Also, more than 80 percent of questionnaire respondents reported a positive impact on their relationships with colleagues and on their self-esteem. Awardees considered the chief benefit of the program to be the improvement in the quality and theoretical grounding of their research and the opportunity to devote large blocks of time to it, an opportunity they were not in the habit of enjoying at their home institutions.

Compared with VPW applicants who had not received awards and women who had received other NSF grants, VPW awardees had authored more published papers following their time as visiting professors and had also more often progressed to institutions with reputations for research excellence. However, because awardees were older and more often held tenured positions, there were fewer gains with respect to tenure and advancement among awardees than among declinees during the period between their awards and the collection of evaluation data.

The evaluators concluded that, for participants, the main benefits were increased research productivity, enhanced professional development, and an expanded network of professional support (see Exhibit 13).

The main benefits of VPW were increased research productivity, enhanced professional development, and an expanded network of professional support.

The evaluation revealed that a disproportionately small fraction of VPW grants had been awarded to women from underrepresented racial and ethnic groups. In response, the NSF evaluation staff recommended that the program be refashioned to better promote diversity within the scientific workforce. In addition, while the evaluation demonstrated that the program had succeeded in expanding research careers of the awardees, NSF felt the impact of the program on the education of young women—through provision of quality teaching, advising, and mentoring—could be improved. This realization contributed to NSF's decision to move the program into the Division of Human Resource Development and consolidate it with NSF's activities for girls and women at the kindergarten through undergraduate levels.

Experimental Program to Stimulate Competitive Research (EPSCoR)

Description. EPSCoR came into being in 1979 in response to national concerns about the geographic distribution of research and development funding. Originally directed at enhancing the capabilities of universities that previously had not competed well for NSF awards, program goals have been broadened to include enhancement of educational opportunities for underrepresented faculty and students. Thus, improvement of science education, coupled with human resource development through the participation of women, minorities, and persons with disabilities, has become central to the program's objectives.

EPSCoR is a partnership program, with NSF funding intended to stimulate local action and investment that will result in lasting

improvements to a state's research infrastructure and increased national research and development competitiveness. The program currently operates in 18 states and in Puerto Rico. Of the 61 institutions involved in the program, the great majority are colleges or universities. During the program's history, NSF has invested more than $120 million, while states have provided more than $275 million in supporting EPSCoR's objectives.

Evaluation. An impact study conducted by Cosmos of Bethesda, Maryland, in 1995 was confined to the educational impacts of EPSCoR and was primarily descriptive. (A comprehensive evaluation of the program is currently under way and due for completion in 1997.) Two categories of program effects were examined—system inputs and outputs. System inputs included the number of enrolled students, outreach activities to recruit targeted student populations, and educational opportunities and experiences offered. System outputs were defined as diversity (proportion and accomplishments of women and minorities) and educational achievements (degrees awarded and other accomplishments of EPSCoR-supported students).

Students have been offered new and revised courses, new degree programs, and access to state-of-the-art equipment as a result of EPSCoR.

Data for the impact study was primarily obtained from existing sources, especially two extensive databases that NSF uses in program planning and evaluation. In addition, information was obtained from annual reports submitted by states and site visits to two award sites (Montana and South Carolina) that were made specifically for the impact study. The evaluation yielded one publication:

- *A Quick Impact Study of the Educational Impacts of the EPSCoR Program (Cosmos 1995).*

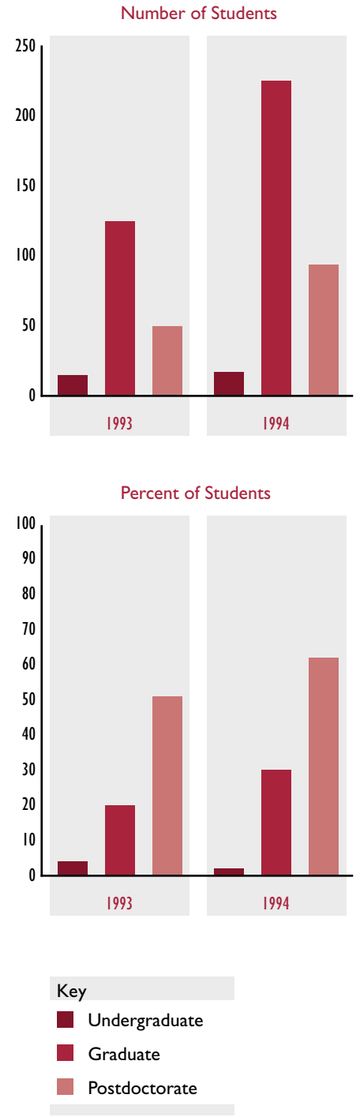
Findings. This study uncovered a wide range of positive educational impacts that may justifiably be attributed to the EPSCoR initiative. The study found, for example, that awardees are supporting either partially or in their entirety a considerable number of diverse outreach activities—ranging from science fair to K-12 teacher training—all of which are aimed at increasing the number of students studying college-level sciences.

EPSCoR funds have also been used to provide direct financial support at all levels of postsecondary education. Evaluators were told that graduate programs have been dramatically strengthened by this support, in many cases raising them to nationally competitive levels.

At each site addressed in the study, students are exposed to a wide range of educational opportunities, including working in research laboratories and copublishing with principal investigators. Additionally, students have been offered new and revised courses, new degree programs, and access to state-of-the-art equipment as a result of EPSCoR. The study also found that EPSCoR-supported students have earned numerous honors and awards, grants and scholarships, and noteworthy placements in graduate schools or the professional work place. High proportions of supported graduate and postdoctoral students have authored or coauthored professional publications (see Exhibit 14).

Moreover, EPSCoR funds have supported women and minorities in two ways: individual support for students participating as part of EPSCoR research teams; and programmatic support for university initiatives, such as conferences, travel grants, and mentoring activities. To varying degrees, these efforts have succeeded. For example, although the proportion of women among EPSCoR-supported personnel in such fields as the physical sciences is higher than the proportion of women involved in these fields nationally, the number of women in EPSCoR-supported engineering and mathematics programs is still quite low. Likewise regarding the participation of minorities, although Asians are well represented, blacks and Hispanics are underrepresented, with the latter two groups constituting about 7 percent of EPSCoR-supported graduate students—about the same as their representation in science nationwide.

Exhibit 14: Number and Percent of Students Participating in Experimental Program to Stimulate Competitive Research (EPSCoR) Who Have Produced One or More Publications, 1993 and 1994



The EPSCoR impact study suggests that the program is strengthening the educational resources and competitive status of the institutions it has funded.

In general, this impact study, documenting a wide range of educational impacts associated with EPSCoR, suggests that the program is strengthening the educational resources and competitive status of the institutions it has funded. However, the evaluators stress the importance of keeping in mind that other programs and activities supported by NSF and by state and private organizations may have contributed to some of the beneficial changes observed in the study.

Although the impact study found that substantial progress had been made by EPSCoR states in improving the participation of women and ethnic minorities, the NSF evaluation staff felt that a greater opportunity was being missed. It recommended that future EPSCoR funding solicitations explicitly address the goal of increasing the diversity of the scientific and technological workforce in these states.

During the past 4 years, the Division of Research, Evaluation and Communication—in addition to completing or initiating an array of studies pertaining specifically to EHR programs—has overseen or coordinated many other activities that support not only its own science and mathematics education assessment efforts but also extend the value of these efforts beyond its own domain.

For example, REC has disseminated a number of documents presenting approaches that recipients of grants from EHR and other agencies might take in developing evaluations, the goal being to elevate the skills of evaluation in general. Also, REC has coordinated several interagency reports and has contributed substantially to the development of various teacher enhancement studies and curriculum surveys that shed light on the extent to which progress is being made in science and mathematics education reform.

REC has disseminated a number of documents presenting approaches that recipients of grants from EHR and other agencies might take in developing evaluations.

Tools for Conducting Evaluations. The evaluations required by EHR's many functions and responsibilities cannot be carried out using a single model. Some *projects* have clear short-term goals—such as increasing the number of minority graduates who obtain bachelor's degrees in science or engineering in one or two colleges or universities—and can be performed using traditional evaluation methods. *Program* goals, on the other hand—such as increasing the interest of women in science careers—are long-term and difficult to measure or attribute to specific program activities. Evaluating changes in educational systems, which requires "systemic evaluation," is an even more complex task. To date, REC has begun to deal with these issues, as is evident in two NSF publications:

- *User-Friendly Handbook for Project Evaluation* was developed to provide principal investigators and project evaluators with a basic understanding of the

evaluation process and the tools to carry out the evaluations required of EHR grantees. These evaluations should be used by project directors and principal investigators to improve their projects as they develop and progress, as well as to document the results that were achieved. Because most projects have measurable short-term goals, the emphasis in the handbook is on quantitative evaluation methods, including sampling, data collection, and report writing. Frequent regional workshops using the handbook are offered to project directors and staff, most of whom have no prior experience in conducting evaluations. A videotape illustrating the handbook's highlights was also prepared. Both the handbook and the workshops have been very well received by the intended audience.

REC has coordinated several interagency reports and has contributed substantially to the development of various teacher enhancement studies and curriculum surveys.

Footprints: Strategies for Non-Traditional Program Evaluation is a collection of commissioned papers authored by experts in the educational research community who had been asked for fresh ideas and new methodologies that might inform the design of EHR evaluations. The papers were first presented and discussed at a conference held under NSF auspices. "Footprints" were defined as evidence of a program's impact on the field of education, on scholarship, on other institutions, and on educational practice. The contributors brought different backgrounds, philosophies, and approaches to the task, and they offered a variety of innovative evaluation schemes. The book offered strong suggestions to adopt mixed evaluation strategies that incorporate both quantitative and qualitative elements.

Innovating and Evaluating Science Education: NSF Evaluation Forums 1992-94 contains selected papers from a series of evaluation forums sponsored by REC for program officers in EHR and other federal agencies with education programs. A paper by Daryl E. Chubin, director of REC, provides the model for EHR program evaluations; the relationship of the design of these evaluations to commitment, burden, and cost; and what NSF expects from program evaluations. Another paper describes how the need for program evaluations became the catalyst for interagency coordination and the development of a master plan for evaluation among federal agencies and funded education programs. One set of presentations focuses on curriculum, pedagogy, and assessment issues. Other papers are concerned with making accessible to underrepresented racial and ethnic minorities (African Americans, Hispanics, and Native Americans) knowledge and skills that come from being learners and planners of scientific processes.

Interagency Support. NSF is only one of several federal agencies that make no substantial investments in science, mathematics, engineering, and technology education. The need for enhanced coordination of federal activities in this area, and increased evaluation efforts by all agencies, has been addressed by several coordinating bodies. In 1993, for example, REC supported an expert review panel and published a report on the panel's findings—titled *The Federal Investment in Science, Mathematics, Engineering and Technology Education: Where Now? What Next?*—which emphasized the need for evaluation of these programs throughout the federal government. REC subsequently coordinated the evaluation of projects supervised by eight agencies that provided opportunities for teachers from around the Nation to study at federal facilities. Participants spent 4 weeks in hands-on education programs in such areas as environmental studies, materials science, and space research. The goal was to improve the science knowledge of the participating teachers, and thus improve middle and secondary school science learning. REC issued a report on this project, titled *Evaluation Report: FCCSET/DOE: 1993 Summer Institutes.*

Currently, a multi-agency study of teacher enhancement and development programs is being conducted by a group of researchers

from Westat, SRI, and the National Center for Improving Science Education. The purpose of this study is to further the understanding of the impact of these projects, starting with those that reflect what are believed to be best practices. A report on the first phase of the project—*Teacher Enhancement/Development Study: A Look at Best Practices: Phase I*—has been completed.

Other Teacher Enhancement Studies. A large portion of the federal investment in science and mathematics education is directed at students in grades K-12, and especially at inservice education for their teachers. Within EHR alone, close to one-third of the fiscal 1995 budget ($200 million out of $600 million) was spent on programs in elementary and secondary education. About half of this amount was allocated to inservice projects to improve the qualifications and effectiveness of mathematics and science teachers. The projects emphasize activities that support NSF's agenda for reform in science and mathematics education. Other government agencies and organizations are also investing significantly in teacher enhancement.

Although a full-impact study of the teacher enhancement is still in planning, REC has published a number of studies that have provided useful background data and perspectives for program managers and policymakers. These completed studies include descriptive information about the activities that the teacher enhancement program has supported in recent years, a broad investigation of issues affecting teacher improvement, and an examination of the characteristics and effects of teacher enhancement activities carried out by public and private agencies in the United States over the past four decades.

Teacher Enhancement Programs: A Perspective on the Last Four Decades, a report prepared by Westat, Inc., is based on extensive bibliographic research. The study included a historical perspective (1950-90) on teacher enhancement programs and examined the major programs currently funded in terms of their goals and impacts. This review concluded that teacher enhancement programs provide significant benefits in terms of several goals: enhancing teacher's acquisition of new knowledge, renewal, and professional leadership. The review pointed out a paucity of studies that assess the impact of such programs on student learning and performance and suggested the need for this focus in the future.

A Study of NSF Teacher Enhancement Programs (TEP) Participants and Principal Investigators: 1984-89 was completed by Abt Associates in 1993. Based on surveys of 450 principal investigators for inservice teacher enhancement projects of the 5-year period and on a sample of 2,000 participants, the survey showed that 63,000 teachers had participated and that teachers felt that their participation had been very beneficial. The main thrust of this study was to document the extent to which several EHR program objectives were met, including introduction of new teaching methods and materials, participation by women and minorities, and participation and additional funding by public and private organizations and through local communities. Most of these program goals were indeed accomplished, according to the self-reports of project staff and participants.

Teacher Enhancement for Elementary and Secondary Science and Mathematics: Status, Issues and Problems. This book, edited by Abt Associates and published by NSF in 1994, consists of chapters written by 11 experts in the field who examined a number of issues, including student learning and curriculum reform, that affect teacher enhancement efforts. The book seeks to help educators who are responsible for planning teacher enhancement programs by sharing what has been learned about successful programs. The major themes are the need for teachers to understand the ways in which students actually learn and the impact of curriculum changes and educational reform on teachers' instructional practices.

An Assessment of the Capacity of Federal Laboratories to Provide Inservice Teacher Enhancement Programs was prepared by Abt Associates and Quantum Research Corporation in 1995. Laboratories of nine federal agencies were surveyed regarding the operation of programs specifically designed to enhance the skills of K-12 teachers and broaden their future capacity to offer such programs. By their own estimates, the federal laboratories already sponsoring teacher

enhancement programs could expand the number of teachers served from the 53,000 involved in 1994 to almost 90,000 teachers served annually. The authors considered the laboratories' estimates to be conservative and consequently used a mathematical model to estimate capacity under alternative scenarios. Many nonparticipating laboratories reported that they could sponsor teacher enhancement activities that would increase the percentage of participating laboratories from 53 percent to 81 percent. The report also provides descriptive information about teacher enhancement programs currently provided and factors influencing the ability of the laboratories to reach maximum numbers of teachers in the future.

Several REC-sponsored surveys have yielded a considerable body of pertinent data on the extent to which school governing bodies and individual educators are responding to the call for improved science and mathematics education.

Surveys of Science and Mathematics Curricula. Additionally, several REC-sponsored surveys have yielded a considerable body of pertinent data on the extent to which school governing bodies and individual educators are responding to the call for improved science and mathematics education. These studies have been far-ranging, covering such issues as the ways in which new scientific knowledge is incorporated into classroom teaching, the attempts being made by teachers to enhance their skills, and the gains that students are showing as a result of reform efforts.

- *A Profile of Science and Mathematics Education in the United States: 1993,* based on surveys conducted by Horizon Research, suggests that much remains to be done to enhance the training of K-12 teachers in science and mathematics. Thus, only 26 percent of elementary teachers feel sufficiently qualified to teach science, and only about one-third of elementary and middle school teachers had spent more than 15 hours on mathematics and science inservice education in the last 3 years.

- *State Curriculum Frameworks in Mathematics and Science: How are They Changing Across the States?,* published by the Council of Chief State School Officers, defines and describes current science and mathematics curriculum frameworks, many of which have undergone major changes in recent years as a result of national, state, and local efforts to improve and reform science and mathematics education. The study examines the extent to which there is consistency across states with respect to curriculum content and the extent to which state standards are consistent with national objectives related to content, skills being taught, classroom implementation, and ethnic and gender equity.

It is hoped that the REC evaluation efforts noted above and others now underway or in the planning stage will put EHR program managers, grantees, and contractors in a better position to achieve the Foundation's goal of improving the Nation's science and mathematics education at all levels.

THE VALUE OF EVALUATION

HIGHLIGHTS

EVALUATION-RELATED ACTIVITIES

FUTURE EVALUATION EFFORTS

FUTURE EVALUATION EFFORTS

In the past 4 years, many aspects of EHR's full evaluation program have been established:

- A regular cycle of evaluations has been created, and evaluations have begun or are in the planning stage for more than half of the Directorate's programs;
- Training in evaluation methods, based on the User-Friendly Handbook for Project Evaluation, was initiated for the Foundation's internal evaluation staff and project directors;
- Interagency leadership and cooperation in evaluation was successfully implemented; and
- A research program supporting evaluation activities was launched.

The major task now facing the evaluation program staff is to complete the evaluations of all EHR programs. Studies about to get underway will examine the following programs:

- Urban Systemic Initiatives
- Rural Systemic Initiatives
- Calculus Reform*
- Graduate Research Traineeships
- Undergraduate Faculty Enhancement
- Advanced Technological Education
- Institution-wide Reform of Undergraduate Education*

* An initiative of the Course and Curriculum Development Program

As the cycle of specific program evaluations approaches completion, REC's evaluation efforts will have greater flexibility and resources to address the broader educational issues associated with several EHR programs, such as teacher preparation and programs for young people.

A good example of such issues is teacher enhancement. Although the effort to improve the scientific and pedagogical skills of the Nation's science and mathematics teachers is the specific focus of one EHR program, related activities play a major role in a variety of other programs, such as the Statewide, Urban, and Rural Systemic Initiatives; Alliances for Minority Participation; and Collaboratives for Excellence in Teacher Preparation. Thus, the subject of teacher enhancement is too broad to be covered in a single evaluation design. Recognizing this fact, REC plans to accumulate data on projects that are oriented specifically toward teacher enhancement; to augment those findings with information gathered from other projects yielding pertinent teacher enhancement evaluation data; and to support new projects that will help fill the gaps in this synthesis of data.

REC's evaluation efforts in the future will have greater flexibility and resources to address the broader educational issues associated with several EHR programs.

Although REC's major focus will continue to be on the evaluation of EHR's programs, the Division's encouragement of better and more frequent evaluations of EHR projects is among its long-term goals. A series of evaluation awareness workshops for NSF staff members and principal investigators is underway. A second series is being planned. It is hoped that these workshops will clarify to NSF staff the Foundation's evaluation expectations.

Also in progress is the development of a second handbook for project evaluators, which will explore the use of quantitative and qualitative models. This handbook is being created in response to requests from current workshop attendees that there be more discussion of how to use qualitative techniques in a rigorous manner.

Several other activities also will contribute to long-term improvement of NSF project evaluations. Through a grant from the American Educational Research Association, a program will be initiated in the fall of 1996 to support selected graduate students with mathematics and science backgrounds in preparing for careers in evaluation. Additionally, REC will encourage the use of stronger evaluation language in NSF program guidelines and will disseminate to prospective grantees good examples of evaluation plans, instruments, and reports.

Moreover, REC will continue to explore opportunities with other agencies.

Evaluation training also is taking a new direction in a joint project between the Foundation and the National Aeronautics and Space Administration (NASA), which is exploring the use of educational performance indicators in evaluating its programs. This effort, emphasizing the management of information in a NASA database, is expected to generate an evaluation training manual and workshop materials that will be useful to federal agencies in addition to NASA.

Moreover, REC will continue to explore opportunities with other agencies as well—for example, a new project that is being carried out in collaboration with the Department of Energy and the Department of Education. This evaluation will look at the impact of federal programs that increase the participation of underrepresented groups in science and mathematics and how these programs can be improved.

Finally, EHR will continue to expand its grants program in evaluation, the goal being to improve evaluation models and techniques and to ensure that evaluations are undertaken of major mathematics and science education initiatives.

ORDER FORM—REPORTS

Copies of NSF Reports are FREE. Check boxes of desired reports and return to NSF one of the three following ways: mail, fax, or e-mail.

See bottom of the back of this form for necessary ordering information.

Elementary and Secondary Education

- ☐ Evaluation of the National Science Foundation's Statewide Systemic Initiatives (SSI) Program: First Year Report, Volume I, Technical Report (SRI International 1994).
- ☐ Evaluation of the National Science Foundation's Statewide Systemic Initiatives (SSI) Program: Second Year Report: Cross-Cutting Themes (SRI International 1995).
- ☐ Evaluation of the National Science Foundation's Statewide Systemic Initiatives (SSI) Program: Second-Year Case Studies: Connecticut, Delaware, and Montana (SRI International 1995).
- ☐ Short-Term Impact Study of the Presidential Awards for Excellence in Science and Mathematics Teaching (Westat 1994).
- ☐ Short-Term Impact Study of the National Science Foundation's Young Scholars Program (Westat 1994).
- ☐ Informal Science Education Program: Evaluation Design Brief and Report of Preliminary Findings (Cosmos 1995).

Higher Education

- ☐ A Short-Term Impact Study of the National Science Foundation's Instrumentation and Laboratory Improvement Program (ILI) (Westat 1996).
- ☐ Evaluation of the Division of Undergraduate Education's Course & Curriculum Development Program, Interim Report (The Network, Inc. 1995).
- ☐ Assessment of the National Science Foundation's 1988-1990 Undergraduate Faculty Enhancement Program: Final Report (Westat 1993).
- ☐ Assessment of the National Science Foundation's 1988-90 Undergraduate Faculty Enhancement Program: Executive Summary (Westat 1993).
- ☐ Assessment of the National Science Foundation's Undergraduate Faculty Enhancement Program: Interpretive Overview. A Statement from the Assessment Advisory Committee (Westat 1993).
- ☐ The Visiting Professorships for Women Program: Lowering the Hurdles for Women in Science and Engineering: Final Report (SRI International 1993).
- ☐ The Visiting Professorships for Women Program: Lowering the Hurdles for Women in Science and Engineering: NSF Summary and Comments (SRI International 1994).
- ☐ A Quick Impact Study of the Educational Impacts of the EPSCoR Program (Cosmos 1995).

Tools for Conducting Evaluations

- ☐ User-Friendly Handbook for Project Evaluation
- ☐ Footprints: Strategies for Non-Traditional Program Evaluation
- ☐ Innovating and Evaluating Science Education: NSF Evaluation Forums 1992-94

Continues on back ▶

ORDER FORM—REPORTS

Interagency Support

- ☐ The Federal Investment in Science, Mathematics, Engineering and Technology Education: Where Now? What Next?
- ☐ Evaluation Report: FCCSET/DOE: 1993 Summer Institutes
- ☐ Teacher Enhancement/Development Study: A Look at Best Practices: Phase I

Other Teacher Enhancement Studies

- ☐ Teacher Enhancement Programs: A Perspective on the Last Four Decades
- ☐ A Study of NSF Teacher Enhancement Programs (TEP) Participants and Principal Investigators: 1984-89
- ☐ Teacher Enhancement for Elementary and Secondary Science and Mathematics: Status, Issues and Problems
- ☐ An Assessment of the Capacity of Federal Laboratories to Provide Inservice Teacher Enhancement Programs

Surveys of Science and Mathematics Curricula

- ☐ A Profile of Science and Mathematics Education in the United States: 1993
- ☐ State Curriculum Frameworks in Mathematics and Science: How are They Changing Across the States?

ORDERING INFORMATION

■ If ordering by mail or fax, please legibly print your name and mailing address in the space provided.

MAIL: NSF
Evaluation Reports Center, Room 855
4201 Wilson Boulevard
Arlington, Virginia 22230

FAX: (703) 306-0434

NAME: _____

ADDRESS: _____

_____

_____

_____

■ If ordering by e-mail, include your name, mailing address, and titles of the reports you are ordering in your message.

E-MAIL: REC@NSF.GOV

The Foundation provides awards for research in the sciences and engineering. The awardee is wholly responsible for the conduct of such research and preparation of the results for publication. The Foundation, therefore, does not assume responsibility for the research findings or their interpretation.

The Foundation welcomes proposals from all qualified scientists and engineers and strongly encourages women, minorities, and persons with disabilities to compete fully in any of the research-related programs described here. In accordance with federal statutes, regulations, and NSF policies, no person on grounds of race, color, age, sex, national origin, or disability shall be excluded from participation in, be denied the benefits of, or be subject to discrimination under any program or activity receiving financial assistance from the National Science Foundation.

Facilitation Awards for Scientists and Engineers with Disabilities (FASED) provide funding for special assistance or equipment to enable persons with disabilities (investigators and other staff, including student research assistants) to work on NSF projects. See the program announcement or contact the program coordinator at:

(703) 306-1636

The National Science Foundation has TDD (Telephonic Device for the Deaf) capability, which enables individuals with hearing impairment to communicate with the Foundation about NSF programs, employment, or general information.

To access NSF TDD dial:

(703) 306-0090

for FIRS dial:

1-800-877-8339

National Science Foundation
Arlington, VA 22230