THE EVALUATION TRAINING PROGRAM: ACCOMPLISHMENTS AND POSSIBILITIES

A REPORT TO THE AERA GRANTS BOARD AND THE NATIONAL SCIENCE FOUNDATION

Jennifer C. Greene

September 2001

Table of Contents

Page

Executive Summary ii

The Context for the Evaluation Training Programs 1

Snapshots of the ETP at Each Institution 3
- Lesley College 3
- University of California at Berkeley 4
- University of Minnesota 5
- Utah State University 6
- A Few Comparative Comments on Program Structure 6

Evaluation Overview and Key Questions 8

Attracting the Right Students 8
- Recruitment Challenges 9
- Recruitment Reflections 10
- Evaluator Comments on ETP Recruitment 12

Offering a High Quality Program to Students 13
- ETP Curricula 13
- ETP Internships 17
- Links to Science and Mathematics 20
- Development of Community and Networking 20
- Evaluator Comments on ETP Design and Implementation 21

Making an Institutional and National Difference 23
- Institutional Outcomes 24
- National Benefits 26

Evaluator Comments Important Considerations for Future Evaluation Training Endeavors 31

References 34

Appendix I: Evaluation Methods 35
Appendix II: Tables 1 and 6 44

List of Tables

Table 1: ETP Fellows: Academic Background and Experience 45
Table 2: ETP Fellows' Perceptions of Program Quality 15
Table 3: ETP Fellows' Perceptions of Program Outcomes 16
Table 4: Dissertation Titles or Topics of ETP Fellows 26
Table 5: Jobs Secured by ETP Graduates 29
Table 6: Jobs and Job Domains Desired by ETP Fellows 49

Executive Summary

In support of other initiatives to improve science and mathematics education in the nation, the National Science Foundation, in collaboration with the American Educational Research Association Grants Program, sponsored an Evaluation Training Program (ETP) from 1997-2001. The ETP sought "to encourage the preparation of evaluators at the doctoral level in order to increase this nation's capacity to competently evaluate its mathematics and science education programs." Four institutions were awarded four-year grants to provide doctoral training in evaluating science/mathematics programs to six students, each of whom received three-year fellowships and research/travel support. The institutions and the number of ETP fellows each supported are: Lesley College (4), University of California at Berkeley (6), University of Minnesota (6), and Utah State University (7).

The institutions were selected to represent a diversity of approaches to doctoral training in the evaluation of science/mathematics education programs. The idea of this initiative was not to replicate a model but rather to learn from diverse institutional and individual experience. While most ETP fellows shared common backgrounds in science/ mathematics and in education (which were program selection requirements), the group of fellows also represented considerable diversity and richness of experience.

Each ETP was required to include coursework in methods and evaluation theory and practice, as well as ongoing field experiences and internships, ideally evaluating NSF projects in science and mathematics education. Towards these ends, Lesley and Minnesota were able to capitalize on their evaluation/research centers which historical housed substantial faculty involvement in science/mathematics education program evaluation. Minnesota also had a tradition of doctoral-level evaluation training. Berkeley faculty also had substantial science/mathematics education program evaluation connections, through various campus centers and institutes (e.g., the Lawrence Hall of Science). These, along with strong assessment and methodological faculty, enabled Berkeley to (re)build their graduate program in evaluation. Utah built on their sound evaluation and methods program, adding new connections to NSF and other federal projects.

Highlights from this summative evaluation of the ETP include the following:

> The majority of ETP fellows experienced strong and vital internships, which provided valued experience in the practical challenges of evaluation and in a science/mathematics substantive domain of major interest to them (for example, curriculum innovation, systemic reform, teacher professional development). The internships were particularly valuable when accompanied by hands-on mentoring by an experienced evaluator and opportunities for reflections on practice.
>
> These ETP curricula would have been strengthened by more methodological diversity and additional coursework or exposure to a wider range of evaluation theories and approaches.

These ETPs successfully developed a sense of community and camaraderie among the fellows, both within and across institutions. Fellows also developed connections to the professional evaluation communities, primarily through fieldwork and conference participation.

Direct connections to NSF were not fostered in this ETP, thus representing a kind of missed opportunity for the agency to strengthen its own evaluation mission and identity.

Overall, the ETP successfully increased our institutional capacity to train evaluators of science and mathematics education programs. Training opportunities exist now that did not prior to the ETP.

And, the ETP successfully prepared a cadre of competent, well trained evaluators of science/mathematics education programs. The vast majority of these ETP graduates intend to include in their career trajectory evaluative work in mathematics and science.

THE EVALUATION TRAINING PROGRAM: ACCOMPLISHMENTS AND POSSIBILITIES

A REPORT TO THE AERA GRANTS BOARD AND THE NATIONAL SCIENCE FOUNDATION

Jennifer C. Greene

September 2001

The Context for the Evaluation Training Program

Substantial improvements in K-12 science and mathematics education remain top national priorities. Not only do children need scientific and mathematical literacy to function as responsible citizens in the future, but our national and global economies demand continued progress in science and mathematics knowledge and its applications (Glenn Commission, 2000). Among the critical domains of mathematics and science education that need improvement are teacher recruitment and preparation, and curriculum innovation and reform.

The National Science Foundation (NSF) has provided important national leadership in mathematics and science education improvements, significantly through the funding of innovative projects. NSF has also provided important leadership in knowledge accumulation about promising educational practices in mathematics and science, primarily through a strong and assertive evaluation requirement for all funded projects. Competent evaluators of mathematics and science educational projects, however, have been hard to find.

Therefore, NSF, in collaboration with the AERA Grants Program, recently sponsored an Evaluation Training Program (ETP), which sought "to encourage the preparation of evaluators at the doctoral level in order to increase this nation's capability to competently evaluate its mathematics and science education programs." The ETP specifically sought to recruit "doctoral students with strong mathematics or science backgrounds and [to provide them with] coursework in evaluation theory and practice, social science research methods, and project-based field training" (Project RFP, April 1996). Through a competitive proposal process, four institutions of higher education were awarded four- year ETP grants (1997-2001) to develop and implement a doctoral program in the evaluation of science and mathematics education programs for students with prior backgrounds and expertise in science or mathematics. Funding in the form of student fellowships and research/travel support was provided for two cohorts of three students

each at each institution — Lesley College, University of California at Berkeley, University of Minnesota, and Utah State University. Under a cost-sharing provision, participating institutions were expected to provide the fellowship funding for two of their six ETP students, one per cohort. Each cohort was supported for three years, and the second cohort began one year after the first. All ETP fellows were to be actively engaged in field-based internships involving evaluations of NSF projects as a critical part of their doctoral preparation.

Support was also provided through the ETP for students and faculty to meet twice annually at the annual meetings of AERA and AEA. In addition to conference participation, these meetings promoted networking among the ETP fellows and faculty through special ETP sessions organized and facilitated by the AERA Grants Board leadership.

In addition to the vitally important policy and program context of science and mathematics education, the ETP engaged the politically charged and intellectually dynamic domain of public program evaluation. During the final years of the 20th century, demands for accountability in our public investments re-surfaced with considerable strength and with support from diverse corners of the political landscape (Shorr, 1998). Advocates, politicians and citizens alike wanted to know not just how many people attended a given program, but what difference that program made in participants' life quality and life chances. It is time, said many, for our social and educational programs to be accountable for discernable and important *outcomes*, for credible and convincing *evidence* that these programs are wise and beneficent investments of our public resources and our public trust. In response to these demands, various faces of program evaluation were summoned. Notably, the GPRA legislation operationalized these demands as requirements for all federal agencies to establish and then meet annual performance standards. Outcomes-based evaluation has been implemented via theory-oriented approaches and logic models. And substantial contemporary legislation, especially in education, requires new programs to be grounded in evidence-based results or programs with demonstrated outcome success.

These significant currents in the evaluation field have all happened within a lively intellectual climate, full of debates about evaluation's public role and responsibilities, its value commitments, and its methods. The liveliness and continuing debates in the field are readily manifested in the recent development and articulation of new evaluation theories (notably, House and Howe, 1999; Mark, Henry, and Julnes, 2000; Schwandt, 1997) and new practices and procedures, for example in mixed-method evaluation (Greene and Caracelli, 1997), participatory

evaluation (Whitmore, 1998), and logic models (United Way, nd; W.K. Kellogg Foundation, 2000).

The ETP thus embodies the intersection of significant ongoing work in science/mathematics/technology education reform and the lively and contested field of program evaluation. This report documents the accomplishments and possibilities of the ETP. The report is intended to supplement prior evaluative documentation and analysis, conducted by members of the AERA Grants Program Governing Board throughout the ETP's implementation, by crystallizing lessons learned about training capable evaluators of science and mathematics education programs and providing important direction for NSF's future evaluator training initiatives.

Snapshots of the ETP at Each Institution

Brief descriptions of the structure and form of the ETP in each participating institution are offered as anchors for the evaluative data to follow. (More detailed descriptions are available in each institution's ETP materials, including the original proposal, recruitment brochures, progress reports, and conference papers.)

Lesley College

At Lesley College, the ETP was located jointly in the Educational Studies doctoral program and in the Program Evaluation and Research Group (PERG), within the Graduate College of Arts and Sciences. The broader doctoral program is a highly individualized, collegial program for returning professionals; the program serves approximately 50 students at any one time. Since its establishment in 1976, PERG has served as a center for educational reform evaluation, including reforms in science and mathematics education and the systemic reforms recently being advanced by NSF. Three faculty provided instructional and programmatic leadership to the ETP at PERG.

ETP fellows at Lesley pursued their own individualized, self-directed program of study, like all Educational Studies doctoral students. Forty-five credits of coursework beyond a Masters degree were required; each student developed a written plan for his/her doctoral studies. This characteristically included three strands: educational theory and research (a year-long interdisciplinary seminar plus varied courses and studies), research and evaluation methodology (at least one quantitative and one qualitative methods course), and evaluation (one course in evaluation and field experiences). The evaluation field experiences for ETP fellows were structured via student internships and other project work at PERG throughout the fellows'

doctoral studies. ETP fellows at Lesley were full members of the PERG community, which provided not only important field opportunities but also valuable support and guidance.

There were four ETP fellows at Lesley College fewer than planned due to recruitment difficulties discussed below. Two fellows began the program in 1997 and two in 1998. All four were women; one is from an under-represented minority group.

University of California at Berkeley

The ETP at Berkeley represented a "newly expanded and revised" programmatic initiative in evaluation within the Quantitative Methods in Education (QME) curriculum. QME is part of a program in the Policy, Organization, Measurement and Evaluation Area of Study within the Berkeley Graduate School of Education. Due to start-up delays, particularly in hiring a new faculty member in evaluation, the Berkeley ETP started one year late, in the fall of 1998. The program originally began as an EdD program but was later changed to include a PhD option, at the urging of the AERA Grants Board.

The ETP curriculum at Berkeley also included three strands: studies in education (department core) and related fields (electives), coursework on evaluation models and methodologies, and evaluation practica and apprenticeships. The department core featured the QME program's traditional strengths in mathematical and statistical methods and in measurement. Practical experience in evaluation was structured as a combination of short-term practica and a third-year apprenticeship, which was envisioned as the groundwork for a dissertation. Field opportunities for ETP fellows, including opportunities in science/mathematics program evaluation, were anticipated through the Lawrence Hall of Science, individual faculty projects, and BEAR (Berkeley Evaluation and Assessment Research), a project initiated in 1994 to serve research and teaching needs in the School of Education. BEAR housed research projects, as well as seminars, visiting speakers, mini-conferences, and other similar activities. Two faculty, one a new hire in 1998, provided most of the vision and leadership for the Berkeley ETP, and two others provided core instruction in assessment and measurement.

Berkeley offered ETP fellowships to six students. Among the first cohort of three students, two began their doctoral work in 1998 and the third had already completed a year of study in QME. Among the second cohort of three students, two began their program in 1999 and the third had already completed two years of study. Four Berkeley fellows are women, two are men. Five are Caucasian, and one is Latino.

University of Minnesota

At the University of Minnesota, the ETP was an interdisciplinary, college-wide initiative in the College of Education and Human Development, supported at the Dean′s level and implemented through the degree-granting authority of three departments: Curriculum and Instruction, Educational Psychology, and Educational Policy and Administration. That is, the ETP fellows could pursue doctoral work in science/mathematics program evaluation through any one of these three departments. (Four fellows chose C&I, and two chose Ed Psych.) The program had an intentional dual focus in both disciplinary training and evaluation. Further, the ETP was formally housed in the Center for Applied Research and Educational Improvement (CAREI). CAREI, which is supported by the College and 40 school districts, is dedicated to forging university-school connections toward meaningful and sustained educational reform. Faculty researchers in CAREI are continuously involved in evaluations of local district educational programs, are nationally known for their work in science and mathematics education research and evaluation, and have strong links to NSF and NSF initiatives. Minnesota is also home to a highly active and visible state affiliate of the American Evaluation Association.

Minnesota ETP fellows completed core coursework in their home department, in research methods (including both qualitative and quantitative methods), and in evaluation theory, methods, and practice. The evaluation practice opportunities were of several types. Students got some fieldwork practice in their courses, and also through their paid employment. The ETP evaluation internship was also required each year. In the first year, fellows worked on some part of an evaluation; in the second year, they planned and conducted a small evaluation from start to finish; and in the third year they worked on a major evaluation component connected to their dissertation. As possible, first and second cohort students were paired in their internships, with the more experienced fellow taking a leadership role. Many of the ETP internships were conducted with NSF projects housed in CAREI. In addition, all fellows were required to participate in an ongoing evaluation colloquium, during which current theoretical issues and problems of practice were discussed with faculty and peers. Two faculty provided the core leadership for the Minnesota ETP.

Minnesota had six ETP fellows. Three students formed the first cohort in 1997, and three the second. When one fellow (Tricia Dieck) left to join her spouse in his new position in Illinois, funding opened up for her replacement. Tricia continued with her graduate work at Minnesota,

although not as an ETP fellow. Five of the six Minnesota fellows are women, and two are African-American.

Utah State University

The ETP at Utah State University was conducted through the Research and Methodology (REM) program, a doctoral program offered since 1985 by the Psychology Department, in collaboration with the Research and Evaluation program in the College of Education. The program requirements for ETP students were very similar to those for other REM students. They included core coursework in research and evaluation methodology (two evaluation courses are part of this core, one introductory and one more advanced), courses in areas of specialized interest, and an evaluation internship. ETP student internships were originally anticipated through the Western Institute for Research and Evaluation (WIRE). WIRE is an off-campus institute run by REM faculty with contracts that are "almost exclusively" in the domains of evaluation and student assessment, many with connections to science and mathematics education, but, upon review, none deemed by the Utah PIs to be of sufficient learning value for ETP internships.

In its proposal, Utah faculty argued for a 4-year ETP at their university, primarily to enable students to "have a significant internship experience," probably a culminating fourth-year experience off-campus, and to take advantage of other opportunities for practical fieldwork that come along during their graduate studies. Utah's proposal indicated they would fund the fourth year for each student. Four Utah faculty have provided program leadership and oversight, one of whom retired during the program; several others have contributed substantially to student coursework.

There were eight ETP fellows identified at Utah State University. Four were recruited in each cohort. But one fellow in the first cohort left graduate school and returned to classroom teaching, leaving seven in the program, five women and two men, all Caucasian. Utah's repeated and extensive efforts to recruit fellows from American Indian groups were, unfortunately, unsuccessful. Utah was able to fund more students because their fellowship rates are considerably less than the funds allocated in the ETP budget.

A Few Comparative Comments on Program Structure

These varied program structures offer several potentially relevant dimensions of comparison. The point of this comparison is not to assess relative worth of the various ETPs because each has

unique value, but rather to signal important components of these pilot endeavors that are possibly related to successful graduate education in science/mathematics education program evaluation.

- *Building on existing strengths — establishing new areas of study:* Minnesota integrated an ETP emphasis into existing programmatic strengths in both evaluation and science/mathematics education. Lesley incorporated the ETP into a long tradition of science/mathematics program evaluation fieldwork, as part of a flexible doctoral program. Utah built on their strengths in evaluation. And Berkeley (re)built their program in evaluation and the connections to science/mathematics anew.
- *Developing evaluation as an area of study:* Minnesota and Utah had established graduate programs in evaluation. Lesley and Berkeley needed to develop these academic programs or areas of study for the ETP.
- *Developing science/mathematics evaluation as an area of practice:* Minnesota and Lesley had established records in science/mathematics program evaluation, including evaluations of NSF programs. The connections to the ETP happened naturally because the ETP was situated *in* the research/evaluation center that conducted science/mathematics program evaluation, and the same faculty were involved with both the ETP and the center. Berkeley had substantial activity in science/mathematics evaluation, but not all located in the same institutional home as their new evaluation program. Thus, connections needed to be forged, and were. Utah′s strong tradition of evaluation included the domains of science and mathematics program evaluation, although not often with federally-funded projects. Thus, connections did have to be developed, and were, with NSF and other federal projects through links with university-based science and mathematics educators and with research firms that conduct federally-funded science and mathematics education program evaluation.
- *Cross-program — within-program location:* Lesley′s program was intentionally interdisciplinary, like its overall doctoral program in Educational Studies. As a college-wide program, Minnesota′s program emphasized both disciplinary education and evaluation. And the programs at Berkeley and Utah were situated both within existing disciplines and within significant cross-disciplinary area and college contexts.
- *Multi-method emphases:* The program at Berkeley was located within a long-standing program that emphasized quantitative methods and assessment. The program at Lesley

fit within the college's tradition of valuing qualitative methodologies. And at Minnesota and Utah, both quantitative and qualitative approaches to social inquiry were emphasized.

Evaluation Overview and Key Questions

The remainder of this report is organized around the key questions that guided this evaluation. These questions are outlined below. The evaluation methods relied heavily on document review, interviews with faculty and fellows, and an ETP fellows questionnaire. Details of the evaluation methodology are found in Appendix I.

The major evaluation questions addressed were:

Attracting the right students:

- How well did the ETP reach its intended audience of people with science and mathematics backgrounds interested in doctoral training in evaluation?

Offering a high quality program to students:

- What was the quality of the program experience for participants?
- What viable alternative program structures were developed?
- In what ways and to what extent did the ETP attain valued outcomes for individual student participants?

Making an institutional and national difference:

- In what ways and to what extent did the ETP attain valued outcomes for participating institutions?
- In what ways and to what extent did the ETP serve to increase the number of competent evaluators of science and mathematics educational programs and to increase diversity in the field of evaluation?

Attracting the Right Students

- *How well did the ETP reach its intended audience of people with science and mathematics backgrounds interested in doctoral training in evaluation? In what ways was the envisioned audience the "right" one?*

The ETP was designed to attract individuals with strong science or mathematics backgrounds so they could substantively understand the programs to be evaluated, as well as backgrounds in education/teaching so they could understand the contexts of the programs to be evaluated.

Specifically, students were sought who had at least a bachelors degree (masters preferred) in mathematics, science, mathematics education, or science education. Teaching experience or coursework in education was also highly desirable. Students could be recruited newly to the institution or drawn from existing graduate student pools at the university.

Table 1 (in Appendix II) presents a listing of the background qualifications of the 23 students who were ETP fellows, as well as the 2 students who withdrew from the program. With respect to students' background in science or mathematics, 20 of the 23 had a bachelors degree in science or mathematics (18) or science/mathematics education (2). Three of these also had a masters degree in science or mathematics. Two of the remaining students had bachelors and masters degrees in psychology/sociology/ educational psychology.[1] The final student had an undergraduate degree in geography and considerable work experience and expertise in technology and assessment. With respect to students' background in education and teaching, 16 had a bachelors and/or masters degree in education, 15 had K-12 teaching experience (most commonly in high school or middle school), and 3 more had teaching experience in a community college or university. The 5 students with no teaching experience brought relevant academic and work experiences in research and evaluation in such fields as nutrition, education, assessment, and disability services.

Recruitment Challenges

Three of the four participating ETP institutions, all except Utah State University, had difficulty in recruiting a substantial applicant pool for the program. This is so even though all institutions used similar advertising and recruitment procedures, notably, announcements in journals and newsletters, postings on listserves, mailings to other universities and to targeted colleagues, and more informal word-of-mouth communications.

The specific recruitment challenges varied somewhat by institution. Lesley College generally attracts older, professional, part-time students who continue to work as they pursue their graduate education. Many of these students simply couldn't give up their jobs to attend school full-time in order to fulfill the ETP timelines and requirements. UC Berkeley planned to recruit students primarily from within the pool of applicants to graduate programs in education and through the identification of promising teachers via evaluation fieldwork. Berkeley's initial inability to attract the right students, combined with a continuing evaluation faculty search, led to

the 1-year postponement of the ETP at Berkeley. Berkeley attributed its recruitment difficulties primarily to the fact that it is a new program in a new area of study (evaluation) for the university and thus not generally well known. In support of this perception, the Berkeley ETP had a smoother second year of recruitment than first year. At Minnesota, recruitment difficulties were attributed to the specific ETP requirements of full-time study and doctoral completion in just three years. These logistical requirements are hard to meet for many people. Minnesota also noted the newness of the program as an additional recruitment challenge. And Minnesota recruited exclusively individuals who were not already graduate students at the University of Minnesota. Minnesota was notably successful in recruiting a highly diverse group of ETP fellows.

As briefly noted earlier, Utah spent concerted efforts in recruiting American Indians for the ETP and did receive 28 applications from American Indians. Unfortunately, with one exception, none of the applicants had the requisite background in science or mathematics. And the one exception could not be enticed to leave her current job to return to school.

Recruitment Reflections

Several of the questions asked on the ETP student questionnaire and faculty interviews related to recruitment issues.[2] In particular, respondents were asked for their views of the importance of the criteria established for the selection of ETP fellows.

Student responses overwhelmingly supported the prerequisite importance of expertise in science/mathematics for competence in evaluating science/mathematics education programs. Twelve (of the 17 respondents to this question) unequivocally viewed expertise in science/mathematics as critical for effective science/mathematics program evaluation. With such expertise, the evaluator can better judge content accuracy and appropriateness and the quality of materials and instruction, said many respondents. Evaluators with content expertise can better understand problems and potential solutions, ask more insightful and probing questions, and better interpret evaluative findings. Such evaluators are also more credible with evaluation participants and funders, noted several respondents. In the eloquent words of one,

---

[1] Early in the program, in the fall of 1998, the AERA Governing Board broadened the student selection criterion of a BA in science or mathematics to include social sciences and elementary education, given significant recruitment challenges in most institutions.

[2] Six of seven faculty PIs were interviewed over the phone. The seventh was unavailable due to travel. A total of 19 of 23 ETP fellows provided information for this evaluation: 16 responded to the ETP fellows questionnaire in writing and 3 via phone interviews. Five of the 16 questionnaire respondents were also interviewed by phone for follow-up purposes.

> *There is a language and way of thinking in mathematics that teachers and curriculum developers are trying to develop in students. If you don't understand this language well enough yourself, it will be impossible to help someone else make progress in furthering student development of the language of mathematics.*

Three additional respondents believed that content expertise isn't essential for evaluator competence, because such expertise is available through other evaluation team members or consultants, as needed. And two respondents, both with backgrounds in social science, did not believe that content expertise was required to be a good evaluator of science/ mathematics educational programs.

Similar questions were asked about the importance of prior expertise in education for competence in evaluating science/mathematics education programs, with highly similar results. Fourteen (of 17 student respondents) believed strongly that first-hand experience in educational settings is vitally important for effective evaluation. Such experience, observed respondents, is especially critical in enabling evaluators to understand the constraints, relationships, and politics of educational evaluation contexts.

> *Schools are highly political, complex environments. Unless a person has been in the trenches' there is no way for [him/her] to understand how different dynamics contribute to implementing curriculum reform.*

Such contextual understanding enables better evaluation practice decisions.

> *It helps the evaluator have a sensitivity to the everyday struggles classroom teachers face in time constraints, motivating the unmotivated students, limited resources, etc. This improves the quality of the questions asked and the methods used to obtain answers.*

Such experience also facilitates the evaluator's work and enhances the evaluator's credibility and legitimacy.

> *My many years of education experience provided me with an emic' or insider's perspective into how schools function. Practice-based experiences in education are simply invaluable in understanding how to prepare for school-site visits and negotiate through and sometimes around possible barriers to your work.*

> *An understanding of the challenges faced by those implementing new programs, along with a background in educational policy and practice at the appropriate school level, are vital ingredients for an evaluation of an educational program. Evaluators can draw on this background to help develop trust, a cooperative spirit, and appreciation for the evaluation process itself among the stakeholders and program participants. It is too easy to dismiss evaluation findings when the evaluator is not perceived as one who understands the environment in which the program took place.*

One respondent again believed that educational expertise could be available from team members or consultants, as needed. And two respondents did not believe educational experience is requisite for competence in science/mathematics educational program evaluation (the same two who did not believe content expertise is required).

Faculty PI views about student recruitment were similar, with a few additional nuances. Faculty generally agreed, although somewhat less unequivocally, that background expertise and experience in science/mathematics *and* in education are important prerequisites for developing competence in evaluating science/mathematics education programs. Several noted that expertise is often partnered with bias, and thus that people without content or context expertise can offer different, potentially valuable lenses. Three faculty noted that the most sophisticated ETP students, those who brought the most critical questions about evaluation, had come to evaluation through "problems of practice."[3] They had experienced first-hand the need for and the potential value of good evaluation. As one respondent noted, professors of evaluation tend to be "methodologists gone practical," and the same groundedness in the continuing practical challenges of teaching and learning is desired in students. Another respondent noted that "professional self-awareness is critical in the evaluation field." Students who have this valuable experience and self-awareness tend to be older and less flexible in their lives, suggesting the need for a program with similar flexibility and options.

No institutional differences in these recruitment reflections were discerned.

Evaluator Comments on ETP Recruitment

Missing from these reflections on what makes a good evaluator of science/mathematics education programs is some consideration of evaluation purpose. Content expertise and contextual experience may be vitally important in evaluations conducted for purposes of *knowledge generation* — learning more about the curricula and instructional processes of effective science/mathematics programs, or for purposes of *program or organizational improvement* (school reform). Different kinds of expertise may be alternatively or additionally useful in evaluations conducted for purposes of *monitoring and accountability* or for purposes of *social critique, democratization, or social change*. Questions of evaluator competence invoke the broader debates about evaluation's purpose and role in society.

[3] All quotes from faculty interviews are verbatim or close paraphrases, as these interviews were not taped.

Offering a High Quality Program to Students

- *What was the quality of the program experience for participants, in terms of such program dimensions as curriculum and course content, pedagogy, mentoring, practical and relevant field experience, networking, and the overall coherence and integrity of the program experience?*
- *What viable alternative program structures were developed? What are key elements of viable alternative program structures?*
- *In what ways and to what extent did the ETP attain valued outcomes for student participants, including:*
  - *individual student mastery of relevant knowledge and skills for entry-level evaluation-related jobs, including critical thinking about evaluation,*
  - *student socialization into a professional evaluation community, and*
  - *membership in a professional supportive network (for students and faculty)?*

In this section of this report, information on program implementation and individual student outcomes related to the following dimensions of ETPs will be presented: curriculum structure, practical evaluation experience, links to science/mathematics, and development of community and networking. These data come from responses to the structured items on the fellows questionnaire, written comments on the questionnaire, and interview responses from both fellows and faculty PIs. Responses to relevant questions from the ETP fellows questionnaire are presented in Table 2 for program implementation and Table 3 for student outcomes. These tabulated responses are presented only for the whole set of respondents, as there was high response variability within institution and no consistent between-institution patterns of difference in these questionnaire data.

ETP Curricula

As described above, all ETP curricula featured coursework in research/evaluation methods and in evaluation models or theories, in addition to core requirements and student electives. From the outset, Lesley and Minnesota planned evaluation seminars for all ETP students and faculty to gather on a biweekly or monthly basis to share work, discuss readings, or trouble shoot problems of practice. Berkeley added an ETP seminar at the request of the Board. Utah's REM program had a required weekly brown bag seminar, which ETP fellows were expected to attend. The internship/ apprenticeship component of ETPs is discussed separately in the next section.

Student reviews of their curricula were generally positive. Most respondents (three-fourths or more) expressed satisfaction with the coursework of their ETP (Table 2, questions 5 and 6). Two-thirds of the respondents rated their progress or accomplishments in quantitative methods as "adequate or excellent," and all but two rated their progress or accomplishments in qualitative methods the same (Table 3, questions 19 and 20).[4] All but one respondent rated their mastery of evaluation theory as "adequate or excellent," and two-thirds gave similar ratings to their understanding of diverse evaluation approaches (Table 3, questions 21 and 22). Yet, when asked in interviews about their own evaluation approach at this point in time, half of the eight fellows interviewed responded by focusing on substantive questions (more research than evaluation) or on existing program goals or funders' questions. One of these commented, "I'm not sure I have a good repertoire of evaluation approaches to choose from." The other four interview respondents offered answers indicative of considered reflection on alternative theoretical evaluation models and/or on challenging issues in practice, for example, the boundaries between program consultation and evaluation. When asked for their views on what evaluation can contribute to high quality science/mathematics education programs, nearly all fellows interviewed responded with some version of "program or organizational improvement," about half also mentioned some form of "accountability," and two mentioned the contributions of evaluation to ongoing, sustained reflective practice.[5]

In terms of curricular improvement, students at Lesley wished for more courses in evaluation theory, students at Berkeley desired more instruction in qualitative methods, and students at Utah desired more evaluation courses overall and more courses specific to science and mathematics education evaluation. But, overall, most students expressed satisfaction with their coursework.

---

[4] Lesley students consistently rated their quantitative skills as weak and their qualitative skills as strong. This is consistent with their curriculum. Students in other institutions may have had less exposure to qualitative methods, and their ratings may reflect more naivet about qualitative skills and accomplishments.

[5] Available syllabi for EPT evaluation theory courses displayed some range of theorists and models, but consistently excluded the more explicitly valued-based models (House, Whitmore, Schwandt) and tended to favor the practical, pragmatic, utilization-focused models (Patton, Alkin, Worthen, Stufflebeam).

Table 2
ETP Fellows' Perceptions of Program Quality (Percents of 17 respondents)*

Question	**SD**	**D**	**U**	**A**	**SA**
5. My expectations regarding what knowledge and skills I would develop as an ETP fellow have been met (so far).			6	76	18
6. My program has/had appropriate coursework requirements.		12	12	47	29
7. My ETP experience has facilitated the development of a sense of community within the ETP fellows group.		12	12	47	29
8. One of the highlights of the program for me has been the connections I have made with NSF personnel and projects.	6	35	18	29	12
9. Once I complete the ETP, I expect to maintain regular contact with some of the evaluation students, evaluation faculty, or NSF personnel I have met during the program.		6	12	29	53
10. As an ETP fellow, I have/had sufficient interactions and engagement with scholars in my mathematics/science domain of interest.		24	18	53	6
11. As an ETP fellow, I have/had sufficient interactions with scientists/mathematicians *about evaluating educational programs* and student learning therein.		40	47	7	7
12. As an ETP fellow, I have/had sufficient interactions and engagement with scholars in education and with practical educational settings.		18	12	59	12
13. As an ETP fellow, I have/had adequate opportunities to practice evaluation in the field.		18	24	29	29
14. As an ETP fellow, I have/had adequate opportunities to practice evaluation in field settings of particular interest and relevance to me.	6	6	24	35	29

* Reported are percents of those responding along a scale of SD=Strongly disagree, D=Disagree, U=Unsure, A=Agree, and SA=Strongly Agree. For question #11, there were only 15 responses.

Faculty PI reflections on curricula and program structure were also mostly positive. One common theme was the breadth and depth of the methodological training ETP students received. Another was that all ETPs were able to build effectively on existing graduate training strengths. Berkeley faculty, for example, pointed with pride to the blend of evaluation and assessment training they had forged in their program, building on their traditional strengths in assessment. And Minnesota faculty highly valued the multiple pathways for graduate study in evaluation they already had in place. Lesley
faculty supported their students' views that more courses in evaluation theories and methodologies were needed, even though this countered the Lesley tradition of self-

Table 3
ETP Fellows' Perceptions of Program Outcomes (Percents of 16 respondents)*

Question: Please assess your progress or end-of-program status on important ETP outcomes.	**Inadeq**	**MinOK**	**Adeq**	**Excel**
16. Familiarity with contemporary science/mathematics education programs.		19	44	38
17. First-hand experience with NSF science-mathematics education programs.	6	19	32	44
18. Expertise in science/mathematics appropriate to the evaluation of science/mathematics education programs.		19	44	38
19. Competence in quantitative applied research and evaluation methods.	7	27	33	33
20. Competence in qualitative applied research and evaluation methods.	6	6	50	38
21. Expertise in evaluation theory and concepts.		6	81	13
22. Understanding of the diversity of extant evaluation approaches and communities.		32	56	13
23. Comfort in selecting an evaluation approach for yourself.		19	44	38
24. Competence in evaluation practice.		19	50	32
25. Practical experience in evaluation.		6	63	32
26. Development of relationships and connections within the professional evaluation community.	6	25	56	13
27. Development of relationships and connections with NSF personnel and projects.	33	40	20	7

* Reported are percents of those responding along a scale of Inadeq=Inadequate, MinOK=Minimally OK, Adeq=Adequate, Excel=Excellent. For questions #19 and #27, there were only 15 responses.

directed study. Utah faculty had endeavored to respond to expressed student desires for courses specific to mathematics/science evaluation, but upon consultation with other experts, affirmed that their existing evaluation courses already addressed key concerns and issues in mathematics/science program evaluation. Lesley and Minnesota faculty underscored the value of situating the ETP within their research/evaluation centers (PERG and CAREI, respectively). With this location, students had multiple and diverse opportunities for practical evaluation experience. Berkeley faculty similarly believed a major strength of their program was student opportunity to participate in "cutting edge" research. Minnesota faculty further underscored the importance of a curriculum that advanced a variety of evaluation models and approaches, implemented by multiple faculty, each with her/his own evaluation approach and practice. The idea is that exposure to diverse evaluation practices can significantly enhance student learning. Multiple faculty are also needed to share the advising load and the many other demands that

evaluation faculty encounter. Berkeley faculty, along with several students from other institutions, agreed that more than one specialist in evaluation is needed.

ETP Internships

Practical field experience in evaluation, particularly evaluation of science and mathematics education programs and especially NSF-funded programs, was a central component of the ETP vision and program design. The majority of students conveyed positive perceptions of their evaluation field experiences, although a significant fraction were less satisfied (Table 2, questions 13 and 14). Even so, over 80 percent of the questionnaire respondents perceived they had attained "adequate or excellent" competence in evaluation practice (Table 3, questions 23-25).

Clearly, practical field experience was a major strength of the ETP at Lesley, significantly attributable to the program's location in PERG. In some important ways, Lesley's collegial apprenticeship model is exemplary. In the words of one student, "my training consisted mostly of opportunities to practice evaluation in field settings." Said another, of great value in Lesley's ETP was "having a window into the evaluation profession becoming engaged in the daily logistics inherent in conducting evaluation work, the challenges of settings, coordinating resources with clients, establishing timetables, etc. — the nuts and bolts!" And a third said, "I am interested in local/regional system math-sci education reform and have had an opportunity to work on projects with such scopes."

Key aspects of the ETP field experience at Lesley were the commitments to ongoing field experience and to student participation in the evaluation process from inception to reporting and utilization. Some students in other institutions lamented that their field experience did *not* include working on an evaluation "from stem to stern." Furthermore, above and beyond their individualized doctoral program, Lesley faculty underscored the important contributions of the PERG culture to the success of their apprenticeship model for the ETP.[6] PERG is an interdisciplinary "community of practice," committed to the use of evaluation in the service of educational improvement and to a collaborative form of evaluation practice in which PERG evaluators are likened to "critical friends." Within these commitments a diverse set of evaluation projects are conducted, including both formative and summative projects with both

[6] It should be recalled that Lesley students are mature professionals and that Lesley's doctoral program rests on an adult learning commitment to self-directed study, factors that also contribute importantly to the success of this apprenticeship model.

quantitative and qualitative designs. At PERG, the culture is one of learning from practice, through reflection, dialogue, and voice. Students are viewed as peers and colleagues, and are expected to fully and actively participate in the ongoing PERG dialogues. ("People might die from dialogue overload," quipped one faculty member.) Student ETP apprenticeships thus were "not private experiences but rather interactive, dialogic engagements with other students and faculty." And the development of ETP fellows' evaluation expertise and identity was thereby nurtured and nourished within the broader collegial self-reflective and critical PERG community of practice. (See also Davis, 2000.)

At UC Berkeley, the ETP also included ongoing opportunities for fieldwork in an intentional interplay of theory and practice. Many Berkeley ETP internships were readily identified in connection with projects in BEAR and at the Lawrence Hall of Science, many of them NSF projects. The substantial level of faculty activity in science and mathematics education at Berkeley contributed to the many high quality internship opportunities for ETP students. Berkeley students reported generally positive experiences in their evaluation fieldwork and internships. Two respondents were very enthusiastic about the ranges of experiences and professionalism they experienced in their fieldwork. Said one, "I am working with Lawrence Hall of Science personnel as an evaluation consultant this year and have worked on a project involving science curriculum for the past three years." Two others expressed the desire for more practical evaluation experience, for experience in a greater variety of evaluation contexts, or for more mentoring and guidance in their practical experience, "more opportunities to contribute and watch expert' evaluators do their work in different settings and contexts. [Just having] more opportunities to do evaluations does not help me learn how to do evaluations."

Like Lesley College, the ETP at Minnesota was institutionally located in their research/evaluation center, CAREI. Like PERG, CAREI is home to a variety of science/mathematics education program evaluations, including many NSF-funded projects. And Minnesota ETP students were able to actively participate in these projects during their three years of evaluation internships.[7] That is, Lesley, Berkeley, and Minnesota all structured evaluation internships as ongoing parts of the ETP curriculum, with increasing responsibilities each year. Student perceptions of their internship experiences at Minnesota were mostly quite positive. "I feel well prepared as a beginning evaluator," and "there were opportunities each

[7] As described above, at Minnesota students also obtained field experience through course projects and through paid research assistant work on research and evaluation projects.

year for me to see evaluators at work and to join in with their efforts" were typical comments from Minnesota students. Three students reported significant benefits from their internship experiences, especially when the projects were closely related to their primary interests (in systemic reform or innovative science curricula, for example). Two others appreciated the multiple opportunities they had but were concerned that they still didn't have enough experience to "take the reins of an evaluation by myself" or to conduct an entire evaluation project, beginning to end.

Finally, the Utah ETP originally planned the internship as a culminating fourth-year experience. And, unlike the other ETP institutions, Utah did not have extensive connections with NSF and other federal projects at the outset of the ETP. At the urging of the AERA Governing Board, Utah worked hard to locate NSF internship opportunities for their ETP students and to incorporate internship experiences on NSF mathematics and science education evaluation projects earlier into their ETP curriculum. Utah student responses to these internship experiences were mixed. Three Utah student respondents were quite positive about their internship experiences. Sample comments include, "the one year internship experience provided the opportunity to learn how to conduct site visits, qualitative interviews and focus groups, and generation of data analysis and final report" and "the internship was also very valuable, as it is difficult to develop a full appreciation for evaluation entirely in the classroom." Two of these students further commented that they wished for more field experience ("more exposure to a variety of potential real-world evaluation problems") and, in particular, for the opportunity "to see an evaluation project from beginning (proposal writing) to end (final report)." Two students reported more negative reactions to their ETP field experiences. Both reported that they were encouraged to work on WIRE projects, but few of these were related to science or mathematics education or to NSF initiatives. These students felt that they did not have the meaningful practical field experience in evaluating mathematics and science education programs that they desired, nor did they have sufficient hands-on guidance and mentoring from an experienced evaluator while in the field. These students stated that their internships minimally engaged science/mathematics evaluation contexts, skills, or professional development.

<u>Links to Science and Mathematics</u>

The grand vision of the ETP included the development or strengthening of links between evaluation and the disciplines of science and mathematics, as well as science/mathematics

education. Some of these linkages were envisioned through ETP student participation in evaluations of NSF projects.

This dimension of the ETP was not realized as envisioned. The individual student experience sometimes included contact with scientists/mathematicians in general or as specific to evaluation (Table 2, questions 10 and 11), but this did not appear to be a salient aspect of the ETP for most students. Many either did not comment on these questions or reported satisfaction with the content knowledge they brought to the ETP. A few students reported that they had strengthened their mathematics/science background while in the ETP — on their own initiative ("it's what *you* make it," said one), through conference attendance, or via their project work. And a few others expressed a desire for more contact or coursework in their discipline, but said these opportunities were not provided or time did not permit them. ETP fellows conveyed somewhat greater satisfaction with the program opportunities provided in education, including science and mathematics education in field settings (Table 2, question 12). And three-fourths or more reported "adequate or excellent" experience with and expertise in contemporary science/mathematics education programs (Table 3, questions 16-18). Institutionally, faculty at Berkeley and Minnesota believed that the ETP had forged somewhat stronger linkages between evaluation and science/mathematics education, primarily through the student internships, many of which were connected to the projects of science/mathematics educators.

Development of Community and Networking

Finally, the ETP also endeavored to advance the professional socialization of fellows, into their own ETP community, into the larger community of program evaluators, and into the particular community of NSF science and mathematics program evaluators.

Development of community was largely successful within the ETP and between ETP fellows and the broader community of evaluators (Table 2, questions 7, 9; Table 3, question 26), but substantially less successful with NSF (Table 2, question 8; Table 3, question 27). Shared office space, regular evaluation seminars, and, for some, joint project work all contributed to ETP camaraderie. And attendance at AERA and AEA, including the cross-institution meetings held at these conferences, were deemed invaluable in forging linkages to the larger evaluation community and to scholars of national stature. Characteristic student comments in this regard are the following. "The travel funds were a tremendous strength of the program, providing exposure to new ideas and new people." And "extremely valuable was the support for

presentations at AERA and AEA conferences and the opportunity this provided to learn of the work of others in the same field, to network, and learn of future employment opportunities."

While some students were satisfied with their NSF contacts, many were "disappointed" with the very minimal contact they had with the NSF science/ mathematics program evaluation community. Students and faculty alike noted that the NSF role in the ETP was undefined and perhaps underdetermined. Half of the faculty respondents commented that NSF should have more actively "brokered" ETP fellow participation in NSF project evaluations, specifically by (1) developing procedures and guidelines for ETP fellows' participation in NSF project evaluations, and then (2) helping to link ETP fellows to particular NSF evaluations. The guidelines needed to address such issues as payment, support for travel and other direct costs, and need for faculty supervision. Beginning evaluation students cannot be expected to conduct an evaluation on their own; even advanced students may still need faculty supervision. A few students and faculty also wished that this ETP had provided opportunities for students to learn more about NSF and its engagement in science/mathematics education — specifically, funding priorities, educational philosophy, current needs for information, views of evaluation. One or two ETP gatherings at NSF would have been very valuable, commented several participants. In short, NSF's silence in the ETP represents a missed opportunity, said a number of students and faculty. ETP participants also noted that with stronger programmatic ties to NSF, ETP graduates would be that much more likely to pursue NSF-related science/mathematics education program evaluation as a primary career option.[8]

Evaluator Comments on ETP Design and Implementation

The RFP for the ETP stated:

> Philosophically, we seek ETPs that recognize that educational evaluation is a creative art, one of wisely applying evaluation theory, subject-matter knowledge, social science and other systematic inquiry, and practical hands-on experience to understand and evaluate educational programs. Students would come to realize that there are no clear, packaged recipes for carrying out an evaluation. Rather, each evaluation is unique in itself. The evaluator must draw on his/her creative problem- and issue-finding skills, knowledge of evaluation theory and practice, practical skills in conducting evaluations and political acumen. Once found and well conceptualized, the evaluator would apply tough-minded analysis to evaluate the program.

[8] Also discussed was the uneven commitment to evaluation in NSF more broadly. A stronger agency-wide valuing of evaluation is needed before evaluation can play a genuinely important role in NSF priority-setting.

This is a laudable vision that was substantially well realized in the ETPs at Lesley, Berkeley, Minnesota, and Utah, each in its own unique way. The ETP curricular emphases on evaluation theory, inquiry methods, and evaluation practice were highly appropriate. Within the inquiry methods curriculum, a diversity of methodologies should definitely be included, perhaps most importantly the traditions of experimental and quasi-experimental design, survey research, and case studies and interview studies. These are the methodologies most commonly used in evaluation practice. And all evaluation training programs should require introductory-level courses in at least two different methodologies plus substantial in-depth work in at least one. These ETPs appeared to approach such methodological requirements with varying success. With respect to the evaluation theory curriculum, most of these ETPs did need **additional coursework in evaluation theory**, particularly to address the diversity of evaluation models and approaches and the ideological and value bases of evaluation as a social practice. Too little of the discussion surrounding these ETPs extended beyond the technical craft of evaluation to include its demands for "creative problem solving and political acumen."

With respect to the evaluation practice component of the curriculum, **an ETP apprenticeship model in which students have ongoing field experience in evaluation, with increasing depth and scope of responsibility** is a strong and viable model for evaluator training. This model enables students to experience a variety of evaluation settings and tasks and, more importantly, to engage in an iterative interplay of learning through theory and practice. As a social practice, evaluation is consequential only in its enactment. Thus, learning about evaluation must meaningfully engage practice. **Effective apprenticeships, however, require mentoring and guidance, as well as opportunities for critical reflections on practice decisions**. Mentoring is essential for apprentices, yet only some of the practice experiences of ETP fellows had sufficient and appropriate mentoring and guidance.[9] And shared opportunities for critique and reflection, for learning from practice, were offered in some but not all ETPs through evaluation colloquia or seminars or through compelling norms of collegial dialogue.

Structurally, the value of fieldwork and apprenticeships is probably enhanced when an **ETP is housed in an active research/evaluation center,** and ETP students can experience most of their evaluation apprenticeships in center projects. The daily coffee pot conversations, brown bag lunches, project meetings and dialogues, and more formal seminars that characterize life in a

[9] In a study in-process on the bases for evaluators' practice decisions (Greene et al.), a surprising number of evaluators cited their mentor or advisor as the most significant influence on their evaluation practice.

research/evaluation center can all contribute to an ETP fellow's ability to make sense of her/his fieldwork experiences and to benefit meaningfully from them. This common site for working and learning clearly also contributes to ongoing student learning about their domain of specialization (science/mathematics education) and to the esprit de corps and sense of community of an ETP cohort.

As noted above, the desired linkages between the ETPs and NSF project evaluation were not realized in these ETPs. **More leadership, initiative, and specific procedural guidance were needed from NSF.**

In the end, clearly, most ETP fellows gained significantly from their ETP experience. When asked, "what is the most valued outcome (so far) of your participation in the ETP?" students responded:

> *ETP provided me with the resources to pursue and complete my dream of getting a doctoral degree and getting the needed training and experiences to competently "do" science evaluation work.*
>
> *The opportunity to earn an advanced degree in a field I am very interested in, gain experience in that field, and get to know some of the key players who define the field.*
>
> *Without the ETP, I'd probably have an MA in instructional technology and be a computer nerd in a high school somewhere. Now, I feel like I'm in a place where I fit. I've found my niche.*
>
> *Development of a new professional identity.*

Making an Institutional and National Difference

- *In what ways and to what extent did the ETP attain valued outcomes for participating institutions, including:*
  - *important linkages and collaborations with other relevant programs and units within the institution,*
  - *visibility and prestige (of the institution's evaluation training program), both within the institution and with other peer institutions involved in evaluation training,*
  - *meaningful balance of costs and benefits of the program, and*
  - *program sustainability without NSF funding?*
- *In what ways and to what extent did the ETP serve to increase the number of competent evaluators of science and mathematics educational programs and to increase diversity in the field of evaluation? Where do ETP fellows go or aspire to go once they complete the*

program, and how closely are their initial jobs related to science/mathematics, education, and evaluation? And in what ways, if any, do the evaluation climate and norms within NSF[10] affect the potential future contributions of ETP graduates to NSF program evaluation work?

The ETP was envisioned to have important impacts beyond the preparation of a skilled cadre of doctoral-level evaluators of science and mathematics education programs. Specifically, the ETP was envisioned to have impacts on the four home institutions and on the national landscape of science/mathematics education program evaluation. This section of this report addresses these program outcomes.

Institutional Outcomes

Institutional benefits. The ETP has made a small but perhaps significant and long-term difference in the evaluation training programs of most of the host institutions. The ETP has helped Lesley College gain some national exposure and has enhanced the visibility and stature of PERG, notably among its clients and within the Lesley community. In particular, the ETP importantly affirmed and strengthened PERG's apprenticeship tradition for training graduate students. Within Lesley College, the ETP has increased attention to science and mathematics education and has contributed to an increased institutional emphasis on evaluation. Lesley faculty noted that being an ETP institution (with regular visits from Audrey Champagne and with Lesley faculty membership on the ETP Advisory Board) helped initiate some conversations about graduate training in evaluation at Lesley, and now, a new graduate program in program evaluation is under development. Said one faculty member, "having the ETP has gotten me in the [Dean's] door to begin with."

At UC Berkeley, the ETP "catalyzed" and "jump-started" the development of an evaluation concentration within QME. Just before ETP there was no evaluation program at Berkeley. Now there is. In addition, the ETP — through its prestigious fellowships — raised the prominence and visibility of evaluation as a field of study and practice within the institution.

In Minnesota, the ETP helped strengthen links between evaluation studies and C&I, particularly mathematics education, as the evaluation faculty brought their own links to science. The ETP also furthered interest within CAREI in evaluations of mathematics and science education programs and enhanced the visibility of evaluation within the Ed Psych Department.

[10] NSF currently allocates full responsibility for individual project evaluation to its PIs, which might undermine the perceived attractiveness of NSF evaluation work to ETP graduates.

Moreover, the ETP contributed momentum and excellent program recommendations — generated from first-hand experience — to the development of a university-wide minor in evaluation studies. While under development for some time, this university-wide minor received a major push from the ETP.

Finally, at Utah, the ETP has contributed to the joining of once-separate college-level and department-level programs in research and methodology.

Program costs. Almost without exception, ETP faculty and students observed that the multi-year, prestigious AERA/NSF fellowships enabled the attraction of highly capable and interesting students to the field of science/mathematics program evaluation. Half of the faculty PIs further observed, however, that the required in-kind match for fellowship funds for one student per cohort was difficult to obtain, especially with no indirect cost returns to the institution. Finally, this ETP experience also illuminated the funding complexities of student fieldwork. In only rare circumstances can student fieldwork contribute independently to the costs of running a graduate program.

Program sustainability. Evaluation training will continue at all four host institutions, albeit with a less exclusive emphasis on science and mathematics education. With PERG, students at Lesley will continue to be offered fieldwork in evaluation, and with Lesley′s flexible doctoral program and possible programmatic evaluation emphases in the future, Lesley students can continue to have "powerful experiences" in evaluation, in the words of one. UC Berkeley will "certainly" continue their new evaluation program; they are eager to help the program grow and expand. The content domains to be emphasized will match those of the students they are able to attract. This is also the case at Minnesota, where the new university-wide minor in evaluation is likely to attract a diverse range of interests from both within the College of Education and throughout the university. A strengthened science/mathematics emphasis in CAREI is an additional legacy of the Minnesota ETP. In Utah, the evaluation program will continue as before the ETP. Although Utah faculty successfully made connections to NSF and science/mathematics evaluation opportunities for the ETP, they are not likely to sustain these in the absence of outside funding or new faculty with science/mathematics evaluation expertise. "We have many far easier opportunities for evaluation projects," commented one.

National Benefits

The ETP successfully developed the mathematics/science education program evaluation skills of a cadre of students. And the ETP funding generated some potentially significant

enhancements of the evaluation training programs of most of the host institutions. The questions now become, did the ETP have a broader impact on the relevant communities of practice? Did the ETP help leverage and advance the importance of evaluation within these broader communities? These questions will be addressed by reviewing the career trajectories of ETP fellows, along with participant perceptions of broader impacts.

Career trajectories of ETP fellows. A substantial minority of the ETP fellows appear headed for a career with a primary commitment to the evaluation of science/mathematics education programs. A significant majority of ETP fellows intend to incorporate both science/mathematics education *and* evaluation as important strands of their professional work, though perhaps not the central strands. Many of these fellows will be mathematics/ science educators, at the K-12 or post-secondary levels, who will also do evaluation. And some of these fellows will be evaluators, whose work will include mathematics and science education. The following discussion substantiates these inferences about ETP fellows′ likely future jobs.

As one indicator of interests at the beginning of a career, Table 4 lists the dissertation titles or domains of interest for the 19 ETP fellows with identifiable dissertation topics at this time.

Table 4
Dissertation Titles or Topics of ETP Fellows

Institution	**Fellow, year of entry**	**Dissertation Title or Topic**	**Math/ science?**	**Eval?**
Lesley College	S.C., 1997			
	M.L., 1997	Historical analysis of science reforms as experienced by Ramapo Indians, NJ, 1970s-1980s	X	X
	M.D., 1998	*On leave* — not yet at dissertation stage		

Institution	Fellow, year of entry	Dissertation Title or Topic	Math/ science?	Eval?
	E.L., 1998			
UC Berkeley	C.K., 1998	Using a Model of Learner Readiness to Study the Effects of Course Design on Classroom and Online Students		X
	M.L., 1998	A cluster evaluation of six school districts′ programs to link school gardens with nutrition and nutrition education	X	X
	M.T., 1998	Using item response theory to predict students′ help needs in intelligent tutoring systems		Testing
	A.B., 1999	Evaluation of the Science for Education and Public Understanding Project (SEPUP), an NSF funded project	X	X
	C.S., 1999	Understanding the limitations and potential for integrating assessment practices into science teachers′ classroom instruction	X	Assess-ment
	D.S., 1999	Measuring the impact of the NCTM standards on AP calculus enrollment and achievement by Latinos and African Americans	X	X
Univ of Minnesota	D.B., 1997	Will be related to high stakes accountability testing in mathematics and science in K-12 schools	X	Testing
	T.F., 1997	Comparing SIMMS and Traditional Programs in Secondary Mathematics Achievement	X	X
	B.J., 1997	Language of Poverty Strategies: Implemented in the Urban Elementary Classroom	X	
	L.C., 1998	The Effect of the "Connected Mathematics Project" on the Middle School Mathematics Achievement	X	X
	A.M., 1998	Will be related to urban environmental problems and solutions		
	B.L., 2000	Will be related to the relationship of multicultural experiences and attitudes in science problem solving	X	
Utah State Univ	C.E., 1997	Utility and Validity of Electronic Resources for Evaluators		X
	L.M., 1997	An Investigation of the Validity Evidence for the Utah Pre-Algebra Criterion Referenced Test	X	Testing
	N.D., 1997	Internalizing Symptoms in Adolescents: Assessments and Relationship to Self-Concept		

Institution	Fellow, year of entry	Dissertation Title or Topic	Math/ science?	Eval?
	T.B., 1998	Related to the effectiveness of an after-school program on student achievement, school attendance, and disciplinary actions		X
	R.G., 1998	not yet at dissertation stage		
	K.W., 1998	The Impact of Electronic Calculators on Mathematics Achievement	X	
	M.N., 1999	Cross-Sectional Concordance and Longitudinal Risk for Depression in Elderly Spouse Pairs		Stats

Acknowledging that dissertation titles can be misleading, that dissertation topics do not always signal career directions, and that an evaluation study by itself is generally *not* accepted as a doctoral dissertation:

- 4 (21%) of ETP fellows do not appear to be pursuing doctoral research in the domains of either science/mathematics education or evaluation,
- 6 (32%) of the dissertations appear to be related to science or mathematics education but not evaluation,
- 4 (21%) appear to be related to evaluation but in domains other than science or mathematics education, and
- 5 (26%) appear to be related to both science/mathematics education and evaluation.

This diversity is one indicator that ETP graduates are likely to choose career pathways that feature science/mathematics education, evaluation, or both, though perhaps not with equal emphasis and centrality.

Table 5 lists the jobs secured by the seven ETP graduates to date. Interestingly, six of these seven currently hold academic jobs, all but one in mathematics or science education. Six of the seven hold jobs related to science and mathematics education, and all seven are or hope to be engaged with evaluation for at least some portion of their time.

Table 5
Jobs Secured by ETP Graduates

Institution	**Fellow, year of entry**	**Job secured**	**Math/ science?**	**Eval?***
Univ of Minnesota	L.C., 1998	Post-doctoral Fellow, University of Minnesota, Mathematics Education and Evaluation; Instructor, Concordia College	X	X
	B.J., 1997	Assistant Professor, University of Central Florida, Science Education and Program Evaluation	X	X
	T.F., 1997	Assistant Professor, Bemidji State University, Mathematics and Mathematics Education	X	(X)
Utah State Univ	L.M., 1997	C&I Administrator, Davis School District, Utah	X	X
	N.D., 1997	Assistant Professor, Linfield College (Oregon), Math Education	X	(X)
	K.W., 1998	Assistant Professor, Lewis-Clark Community College (Idaho), teaching math/science methods and educational assessment	X	(X)
	M.N., 1999	Assistant Professor, Utah State University, Family and Human Development		(X)

* Graduates who indicated an intention to incorporate evaluation into their work, as a minor or secondary emphasis, are noted with an (X).

ETP fellows were also asked about their career aspirations and specific plans on the fellows questionnaire/interviews. Table 6 (in Appendix II) presents a tabulation of the 18 responses received. Fellows′ specific career hopes and intentions evidenced some variability in desired job location: six wanted a faculty position in a university (4) or community college (2), three wanted to work for a larger research/evaluation center, three wanted to work directly in schools and classrooms, two wanted to do evaluation work in a community location, and one planned to work for a state department of education. (The other three did not specify a job location.) With respect to substantive field, all but one expressed a desire or intention to be involved in science/mathematics education in the future; three-fourths expected active or extensive involvement. And all respondents expressed a desire or intention to be involved in evaluation in the future. About three-fifths perceive evaluation as central to their future work; the remainder see it more as a secondary focus or see themselves more as technical or support staff in an evaluation rather than a PI.

So, yes, the ETP has contributed to the national landscape a group of generally well trained evaluators of science and mathematics education programs, most of whom wish to use these skills in their professional work lives. Moreover, both ETP fellows and faculty credit the ETP for attracting and providing evaluation training for "very good" students, many of whom would have been unlikely to find work opportunities in mathematics/science education program evaluation on their own. Said Berkeley faculty, "all of our students will heavily incorporate evaluation into their careers. And all but one would not have emphasized evaluation very much at all without the ETP." "We have had very good students who normally wouldn't have come back to school to be evaluators. And they will make very good evaluators." Said Utah faculty, most of our students are "real gems," even better than the fine students we usually get in the REM program. And Lesley faculty commented on the significant "cognitive shift" effected by their students' ETP participation and, specifically, learning about evaluation. Through this shift, Lesley students learned about the importance of "remaining in a position of inquiry" as they engage with their own and others' educative practices. These fellows will incorporate this evaluative thinking into whatever work they do in the future.

When asked if their career aspirations had changed since their participation in the ETP, about one-third of the respondents said no, for example, "No. If anything, ETP has confirmed for me that what I enjoy most is teaching and building science ed programs and curricula." Another one-fourth of the respondents said their career aspirations had been broadened and expanded, for example, "now I see several more positions that I can fill" and "I've always been interested in outcomes assessment , but on a focus on evaluation broadens my scope to include issues relative to the broader improvement of educational practice than one classroom or teacher at a time. " And the remaining respondents (close to one-half) said "absolutely" and "yes, I had never previously thought about program evaluation before participating in this program," and "yes I now know about the field of evaluation and the opportunities that exist."

Finally, the ETP did make progress on its intention to increase the diversity of evaluators of science/mathematics education programs. Among the 23 fellows, 18 are women and 4 (3 women, 1 man) are from under-represented minority groups. In the opinion of one faculty PI, increasing the diversity of science/mathematics education program evaluators remains a high national priority, because minorities continue to be under-represented in all domains of science and mathematics, and because science and mathematics are "power areas."

Evaluator Comments: Important Considerations for Future Evaluation Training Endeavors by NSF

Continued efforts by NSF to augment the ranks of competent science/mathematics education program evaluators are clearly warranted. The potential for focused training efforts to accomplish this goal is clearly demonstrated by this ETP. It is understood that less costly alternatives than the institutional model of this ETP are needed in the future. It should be fully acknowledged, however, that **this ETP successfully and importantly enhanced the institutional capacity for training evaluators of science/mathematics programs.** A new evaluation program now exists at UC Berkeley, a university-wide evaluation minor now exists at Minnesota, and increased emphases on science/mathematics evaluation persist at Berkeley, Minnesota, and Lesley. Given this demonstration of successful capacity building, if institutional ETPs were ever to be funded again sometime in the future, there should be little hesitation to consider as viable applicants promising or nascent programs, in addition to existing programs, and a diversity of program structures, as was evident in this set of pilot ETPs.

Any future funding model for evaluation training in science/mathematics education should consider the following program characteristics.

- Given the generally preferred **program length of four years**, ETP fellowships in the future could be offered for three years, with the host institution expected to fund the fourth, either through fellowship funds or through research-teaching assistantships. **Conference travel costs** should also be included in any future ETP fellowships, as the opportunity to experience and participate in national conferences and to build professional relationships was universally highly prized.
- Candidates for evaluation training should have **a commitment to future work in science/mathematics education** and thereby probably past expertise and experience in science or mathematics. In these ETPs, students who came to their graduate-level evaluation training *without* a background in science or mathematics appear to be likely to pursue professional work in evaluation, *but* in domains other than science and mathematics education. Students who came with a background in science or mathematics are all planning to continue to work in these domains and can contribute their evaluative skills and evaluative thinking to their work, even if they pursue evaluation as a secondary professional activity.

- The locations for evaluation training (institutions, institutes, or other) should have most importantly **"good people"** — faculty/instructors/project leaders who can provide the critical mentoring needed in effective graduate evaluation training, especially fieldwork. Good mentors are probably themselves evaluation practitioners with experience evaluating science/mathematics education programs, notably reflective practitioners, so what they can share with students is expertise grounded in experience. For NSF goals, good mentors are also practitioners with connections to the NSF network of science/mathematics education evaluation, so they can share these connections with their students. Good mentors are also committed teachers, willing to spend the time needed in nurturing students' professional development.[11]
- The locations for evaluation training should ideally have **viable and strong "communities of practice" in science/mathematics education program evaluation**, communities that would welcome apprentice trainees into their ranks. In this ETP, there was considerable value in having as a home base for ETP fellows a center with lots of science/mathematics evaluation activity already in action. There was similar although less powerful value in the ongoing ETP evaluation seminar conducted at each host institution. Being a member of an established community of evaluation practice in science and mathematics contributed in many important ways to students' skills acquisition and professional socialization.
- Training should include an **apprenticeship model in which students have ongoing field experience in science/mathematics education program evaluation, with increasing depth and scope of responsibility**. As noted above, evaluation is most importantly a social practice that becomes meaningful only in its enactment. Because evaluation is not simply the application of learned techniques, but rather the artful crafting of unique socio-political responses to unique contexts, practice is essential. Again, an effective apprenticeship model relies on good mentoring.

Any future funding model for evaluator training in science/mathematics education program evaluation can also benefit from a **more substantial commitment by and more significant involvement of NSF**. Many ETP fellows lamented their limited opportunities to connect with NSF projects and personnel. This represents an important missed opportunity for NSF to

[11] The considerable literature on mentoring could also be consulted here for additional, more specific ideas. See, for example, Hegsted (1999). Formal mentoring as a strategy for human resource development: A review of research.

strengthen its own evaluation culture and to importantly attract ETP evaluators to work on NSF projects. In the absence of significant interactions with NSF, fellows are left wondering if there will indeed be NSF-related job opportunities for them upon program completion.[12] A strong program model will help fellows build their NSF evaluation career while in their graduate program. Mechanisms to accomplish this, generated by ETP respondents, include:

- Ensure that most/all training apprenticeships are on NSF projects.
- Provide a "certificate of NSF evaluation training" to all program graduates.
- Provide a list of "NSF-trained evaluators" to all NSF PIs.
- Hold a grants competition for fellows only, to develop evaluation proposals for selected NSF projects. Fund the "winners" of this competition to actually do the NSF evaluations.
- Open up postdoctoral evaluation positions for people with training in evaluating science/mathematics education programs, positions with NSF or with NSF projects in the field.

NSF can also begin to reap even more benefits from its recent investments in evaluation training (the ETP, the *User-Friendly Handbooks*, various short-term training activities, and various workshops for reflection and future planning), but only if these investments can build on and strengthen one another into the future.

Human Resource Development Quarterly, *10* (4), 383-390.

[12] In this regard, the ongoing involvement of the AERA Grants Board members in the ETP must be recognized as a significant contribution to program quality and effectiveness. Without exception, fellows observed that they had benefited from interactions with scholars of such stature as the ETP Governing Board panel — benefits that included both professional development and networking and personal validation.

References

Davis, F. (January 2000). "Training mathematics and science education program evaluators: A perspective from Lesley College." Presentation at NSF Workshop on Training Mathematics and Science Evaluators, Arlington, VA.

Glenn Commission (2000). *Before it's too late: A report to the nation from the National Clearinghouse on Mathematics and Science Teaching for the 21st century*. Washington, DC: National Science Foundation.

Greene, J.C., and Caracelli, V.J. (eds.) (1997). *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms. New Directions for Evaluation no. 74.* San Francisco: Jossey Bass.

Hegsted, C.D. (1999). Formal mentoring as a strategy for human resource development: A review of research. *Human Resource Development Quarterly, 10*(4), 383-390.

House, E.R., and Howe, K. (1999). *Values in evaluation and social research.* Thousand Oaks, CA: Sage.

Mark, M.M., Henry, G.T., and Julnes, G. (2000). *Evaluation: An integrated framework for understanding, guiding, and improving policies and programs.* San Francisco: Jossey-Bass.

Schwandt, T.A. (1997). Evaluation as practical hermeneutics. *Evaluation, 3*(1), 69-84.

Shorr, L.B. (1998). *Common purpose: Strengthening families and neighborhoods to rebuild America.* New York: Doubleday.

United Way of America (nd). *Measuring program outcomes: A practical approach.* Alexandria, VA.

Whitmore, E. (ed.) (1998). *Understanding and practicing participatory evaluation. New Directions for Evaluation no. 80.* San Francisco: Jossey-Bass.

W.K. Kellogg Foundation (2000). *Logic model development guide.* Battle Creek, MI.

Appendix I
Evaluation Design and Methods

Evaluation Context

The ETP was evaluated on an ongoing basis by members of the AERA Grants Program Governing Board. Board members made annual site visits to each ETP institution — where they reviewed documents, met with faculty, and interviewed fellows — and provided active oversight and leadership to the annual meetings of ETP fellows and faculty. For the final year of ETP implementation (2000-2001), a more summative evaluation of the program was enabled by a supplemental one-year SGER grant from the NSF (November 2000 to October 2001).

Evaluation Overview

The primary purpose of this ETP evaluation was to crystallize lessons learned about training competent evaluators of science and mathematics education programs and to provide important direction for NSF's future evaluator training initiatives. Evaluation staff in NSF, the AERA Grants Program Governing Board, and PIs at each participating institution were the key audiences for this evaluation.

The major questions addressed were:

Program implementation:

1. How well did the ETP reach its intended audience of people with science and mathematics backgrounds interested in doctoral training in evaluation? In what ways was the envisioned audience the "right" one?
2. What was the quality of the program experience for participants (students, faculty, important others), in terms of such program dimensions as curriculum and course content, pedagogy, mentoring, practical and relevant field experience, networking, non-academic support (stipends, housing, and so forth), and the overall coherence and integrity of the program experience?
3. What viable alternative program structures were developed? What are key elements of viable alternative program structures?

Program outcomes:

4. In what ways and to what extent did the ETP attain valued outcomes for student participants, including:
   - individual student mastery of relevant knowledge and skills for entry-level evaluation-related jobs, including critical thinking about evaluation
   - student socialization into a professional evaluation community
   - membership in a professional supportive network (for students and faculty)
5. In what ways and to what extent did the ETP attain valued outcomes for participating institutions, including:
   - important linkages and collaborations with other relevant programs and units within the institution
   - visibility and prestige (of the institution's evaluation training program), both within the institution and with other peer institutions involved in evaluation training
   - meaningful balance of costs and benefits of the program

- ♦ program sustainability without NSF funding

6. In what ways and to what extent did the ETP serve to increase the number of competent evaluators of science and mathematics educational programs and to increase diversity in the field of evaluation? Where do ETP fellows go or aspire to go once they complete the program, and how closely are their initial jobs related to science/mathematics, education, and evaluation? And in what ways, if any, do the evaluation climate and norms within NSF[13] affect the potential future contributions of ETP graduates to NSF program evaluation work?

Evaluation Methods

This evaluation used a case study design within an overall programmatic framework. That is, evaluative information from each of the four participating institutions was summarized and synthesized using common program dimensions of quality and effectiveness (primarily those dimensions incorporated into the evaluation questions).

The following methods were used:

1. Existing descriptive and evaluative information on the ETP program structure and implementation and on program progress and quality for each participating institution was reviewed and summarized. This information included for each institution (a) original program proposals, (b) annual progress reports and the final project report, (c) reports from annual site visits by AERA Board members to each institution, (d) AERA Grants Board annual reports, and (e) other relevant information available (for example, correspondence, samples of student work, results of student surveys or interviews, and the like). Much of the needed information on the quality of program implementation (evaluation questions #1-#3) was secured through this document review.

2. All 23 ETP fellows were sent a fellows questionnaire (a copy of which is attached) in April 2001. The questionnaire assessed fellows′ perceptions of recruitment, program quality, and both individual and programmatic outcomes. Fourteen (61%) of the fellows completed this questionnaire upon first request, and one declined to complete it. In early May, the questionnaire was re-sent to the 8 fellows who had not responded. Two fellows completed the questionnaire upon this second request, bringing the total number of questionnaire respondents to 16 (70%).

   Three additional fellows agreed to a phone interview in lieu of completing the questionnaire. For these fellows, a combination of the questionnaire and the fellows interview guide (see copy attached) was used to elicit their ETP program experiences and perceptions. In addition, phone interviews were also conducted in May and June with 5 fellows who had already completed the questionnaire. The purpose of these interviews was to follow-up in more depth questions about program experiences and outcomes.

   Thus there were a **total of 19 fellows (83%) who provided feedback for this evaluation.** Eleven fellows completed only the questionnaire, 5 fellows responded to

[13] NSF currently allocates full responsibility for individual project evaluation to its PIs, which might undermine the perceived attractiveness of NSF evaluation work to ETP graduates.

both the questionnaire and the phone interview, and 3 fellows responded only to the interview (which for them also included a few questionnaire items). The breakdown of the 19 respondents by institution is 3 of 4 from Lesley, 5 of 6 from Berkeley, all 6 from Minnesota, and 5 of 7 from Utah.

3. All ETP PIs were invited to participate in a phone interview, which focused on assessing their perceptions of the quality of the ETP experience and information on important institutional outcomes (question #5) and on potential ETP impact on the field (question #6). Six of the seven PIs were able to participate in this interview, at least one from each of the four institutions.

4. PIs and AERA Grants Board members had a further opportunity to contribute their views of the ETP program at a discussion of a draft of this evaluation report at a meeting in July 2001.

The Evaluation Training Program: Perspectives and Accomplishments
Participant Questionnaire, April 2001

This questionnaire asks for your views on your ETP experiences and achievements. These data are being gathered as part of a final evaluation of the ETP program. Results will be shared with all ETP participants and funders as input to continuing discussions about future training possibilities for evaluators of science and mathematics education programs. Your responses will be aggregated with others at your institution, while maintaining strict confidentiality of individual responses.

*Please complete and return this questionnaire electronically to me (jcgreene@uiuc.edu) by **April 25**. If you prefer to share your responses with me on the phone, please email me to arrange a time to do so at your earliest convenience. Or, you may print the questionnaire, respond on the hard copy and return to me via mail, so that it **arrives by April 25**. (My postal address is on the covering email.)*

Many thanks in advance for your time, your thoughtfulness, and your cooperation.

PART I: RECRUITING AND EDUCATING FUTURE SCIENCE AND MATHEMATICS EDUCATIONAL PROGRAM EVALUATORS

1. Recruiting the right people has been one challenge the ETP has faced in most institutions. Please record the experiences and expertise you brought to and then developed within the ETP, as requested in the following chart.

Domain of relevant exp	Brought to ETP *(what degrees and/or experience)*	Developed or strengthened in ETP *(yes, somewhat, no)*	Wish I had more of this during the ETP *(fill in specifics)*
Science, mathematics			
Education			
Evaluation (theory and practice)			
Other_____			

2. In your view, in what important ways are prior expertise in science/mathematics prerequisite for competence in evaluating science/mathematics education programs?
[Note: Please take as much space as you need for comments on this and all other open-ended questions. Do not feel restricted by the small space allocated, which is only an artifact of the electronic design.]

3. In your view, in what important ways is prior expertise in education prerequisite for competence in evaluating science/mathematics education programs?

For questions 4-13, please use the following key to underline or otherwise mark the one response that comes closest to representing you own views.

SD	=	**Strongly disagree**
D	=	**Disagree**
Unsure	=	**Neither agree nor disagree**
A	=	**Agree**
SA	=	**Strongly agree**

4. I would not have been attracted to the ETP without the offer of a fellowship. **SD D Unsure A SA**

5. My expectations regarding what knowledge and skills I would develop as an ETP fellow have been met (so far). **SD D Unsure A SA**

   What most importantly explains your marked response?

6. My program has/had appropriate coursework requirements. **SD D Unsure A SA**

   What most importantly explains your marked response?

7. My ETP experience has facilitated the development of a sense of community within the ETP fellows group. **SD D Unsure A SA**

8. One of the highlights of the program for me has been the connections I have made with NSF personnel and projects **SD D Unsure A SA**

9. Once I complete the ETP, I expect to maintain regular contact with some of the evaluation students, evaluation faculty, or NSF personnel I have met during the program. **SD D Unsure A SA**

10. As an ETP fellow, I have/had sufficient interactions and engagement with scholars in my mathematics/science domain of interest. **SD D Unsure A SA**

    What most importantly explains your marked response?

11. As an ETP fellow, I have/had sufficient interactions with scientists/mathematicians *about evaluating educational programs* and student learning therein. **SD D Unsure A SA**

    If "agree or strongly agree," what specifically was the content of these interactions?

12. As an ETP fellow, I have/had sufficient interactions and engagement with scholars in education and with practical educational settings. **SD D Unsure A SA**

    What most importantly explains your marked response?

13. As an ETP fellow, I have/had adequate opportunities to practice evaluation in the field. **SD D Unsure A SA**

14. As an ETP fellow, I have/had adequate opportunities to practice evaluation in field settings of particular interest and relevance to me. **SD D Unsure A SA**

    Please further explain your responses to 13 and 14.

15. What else about the design and implementation of the ETP at your institution has mattered to you? It is especially important to understand which program components, dimensions, or experiences are/were particularly meaningful or helpful to you, and which are/were of little value or importance for you.

PART II: ENABLING INDIVIDUAL ACCOMPLISHMENTS

Please assess your progress or end-of-program-status on important ETP outcomes, using the key that follows. Underline or otherwise mark the one response that comes closest to representing your views.

Inadeq	=	**Inadequate**
MinOK	=	**Minimally OK**
Adeq	=	**Adequate**
Excel	=	**Excellent**

In the final column, underline or mark the checkmark (ˆ) if you can attribute this outcome mostly to your participation in the ETP.

16. Familiarity with contemporary science/mathematics education programs **Inadeq MinOK Adeq Excel ˆ**

17. First-hand experience with NSF science/mathematics education programs **Inadeq MinOK Adeq Excel ˆ**

18. Expertise in science/mathematics appropriate to the evaluation of science/mathematics education programs **Inadeq MinOK Adeq Excel ˆ**

19. Competence in quantitative applied research and evaluation methods **Inadeq MinOK Adeq Excel ˆ**

20. Competence in qualitative applied research and evaluation methods **Inadeq MinOK Adeq Excel ˆ**

21. Expertise in evaluation theory and concepts **Inadeq MinOK Adeq Excel ˆ**
22. Understanding of the diversity of extant evaluation approaches and communities **Inadeq MinOK Adeq Excel ˆ**
23. Comfort in selecting an evaluation approach for yourself **Inadeq MinOK Adeq Excel ˆ**
24. Competence in evaluation practice **Inadeq MinOK Adeq Excel ˆ**
25. Practical experience in evaluation **Inadeq MinOK Adeq Excel ˆ**
26. Development of relationships and connections within the professional evaluation community **Inadeq MinOK Adeq Excel ˆ**
27. Development of relationships and connections with NSF personnel and projects **Inadeq MinOK Adeq Excel ˆ**
28. Other _____ **Inadeq MinOK Adeq Excel ˆ**
29. For you, what is the most valued outcome (so far) of your participation in the ETP?
30. At this point in time, what do you hope/plan to be doing or anticipate doing after you complete the ETP?

    To what degree do you hope or expect to be actively involved in mathematics/science education programs?

    To what degree do you hope or expect to be actively involved in evaluation?

    Have your future job/career aspirations changed since you have been involved in the ETP? If not, why not? If so, in what directions and for what reasons?

PART III: SUSTAINING PROGRAM ACCOMPLISHMENTS

31. Do you expect your institution to maintain a doctoral program in evaluation of science/mathematics education programs?

    What in particular is most relevant to this question of sustainability?

32. Do you expect to continue your own professional development in the evaluation of mathematics/science education programs?

    If so, to what sources would you turn first for this professional development?

33. What should have been asked on this questionnaire, but wasn′t? Please add important further ETP perspectives and accomplishments for you.

Would you be willing to participate in a follow-up interview? ***YES*** ***NO***

THANK YOU SO VERY MUCH FOR YOUR TIME AND THOUGHTFUL RESPONSES!

ETP, FELLOWS INTERVIEW GUIDE

Greatly appreciate your taking the time, especially since you completed the questionnaire and have also given this kind of feedback throughout your ETP career. From our conversation, I will not attribute any comments you made to you individually, but I would like to aggregate responses by institution, because the programs have differed in important ways at the four institutions. I hope that will be OK with you? If any comments arise that you wish to make more confidentially, just say so and I will fully honor your request.

1. I would like to begin with some discussion of the "center", which was a component of the ETP at _____. Could you first describe for me the nature of the ETP's connections to or relationship with the "center"?

   Thanks, now how vital do you believe these connections were to the quality of your ETP experience? Or, what would your experience have been like without these connections to the "center"? And how might these connections have been improved?

2. I would like to shift now to your evaluation perspectives at this time. Could you please describe your current views on how best to evaluate science/mathematics educational programs in the public schools?

   [As needed, envision an innovative NSF science/math middle school program being piloted in a medium-sized urban school district. Program features problem-based curriculum, project-based and active student learning, as well as some creative technological tools for teaching and for student work.]

   Ask first for general description and then probe each of these sub-areas.
   - purpose and audience
   - key evaluation questions
   - criteria for judging quality
   - design and methods
   - uses

   What most importantly influenced these evaluation ideas?

3. Could you now share with me your current views on the kinds of contributions evaluation can make to mathematics and science education?
   What are the important roles for evaluation in the broad challenge of improving our public education system in mathematics and science?

   And what most importantly influenced these views?

4. Please now talk about your current career aspirations. What do you hope to do upon completion [refer to questionnaire responses]?

   Has the ETP influenced your thinking about your future career [refer to questionnaire responses]? If so, in what ways?

5. What recommendations do you have for future ETPs?

6. Finally, what else is on your mind that might be helpful to me and others concerned about the evaluation of science and mathematics education programs?

Thank you!

ETP, Interview Guide for PIs

I appreciate your taking this time, especially since you have done this many times before. I hope to be succinct and focused. And while I will not attribute any general comments about the ETP to you specifically, it will be helpful to be able to identify the more specific reflections you have about the ETP to the program at your institution. I hope that will be OK with you? If any comments arise that you wish to make confidentially, just say so and I will fully honor your request.

1. I would like to begin with two questions about the design and implementation of the ETP. First is the knotty issue of student recruitment. At this point, after your several years of experience, could you please share your views on the particular characteristics and academic and experiential backgrounds that are *most important* for people entering a science/mathematics evaluation graduate program? In other words, in what ways did the ETP look for the right kinds of students?

2. Now, also from your vantage point of several years of experience, what would you say now are the one or two *major strengths* of your evaluation training program? And what are the one or two features of your program you would *do differently* the next time around?

Now, for the rest of our discussion, I would like to turn to *what difference the ETP has made* — for individual students, for your institution, and for the broader community of science/mathematics educational program evaluators.

3. To the best of your knowledge at this time, what will your ETP fellows be doing once they complete the program?

   How satisfied or fulfilled are you with these job/career choices? Is this what you hoped would happen with your ETP graduates?

   More broadly, what do you consider to be important and useful careers for ETP graduates?

4. Now, I would like to turn to your reflections about the institutional location, benefits, and longer-term continuation of the ETP at ___?

   Let′s start with your thoughts about the advantages and limitations of locating the ETP within ___? Specifically, did this institutional location enable relevant linkages with content experts in science and mathematics? With science and mathematics educators? With other relevant experts and resources (identify)?

   Next, were there any benefits from ETP for your program or your domain of work within your overall institution? Such as visibility, stature, resources?
   Were there any benefits from ETP for your overall institution?

   And, now what you expect to happen with the ETP at your institution in the future? Please elaborate.

   Summary and check.

5. The last tier of outcomes to discuss relate to what difference the ETP has made in the broader community of evaluators of science and mathematics educational programs. In your view, what *has* the ETP contributed to this broader community? Specifically, to the visibility, quality and effectiveness of math/science educational program evaluation?

6. And further, what recommendations do you have for any future ETPs?
   Under what conditions would you bid on another RFP?

   Summary and check.

7. And finally, what else is on your mind that I haven′t yet addressed?

Thank you!

Appendix II

TABLES 1 AND 6

Table 1
ETP Fellows: Academic Background and Experience

INSTITUTION	Fellow, year of ETP entry	Science/math background	Education background	Other relevant background
Lesley College	S.C., 1997	BS biology	BS secondary education MA in special needs education Science teacher	Earthwatch Expedition Science consultant with public broadcasting station Volunteer work with scientists in the field Prior graduate coursework in evaluation
	M.L., 1997	BA chemistry	MAT physical science Secondary science teacher	Prior PERG evaluation work
	M.D., 1998 *—On leave*	BS physics	MAT physics teaching Science teacher K-12 Science coordinator, K-12	Prior work in industry Associate director of Science Northeastern Field test coordinator for Active Physics
	E.L., 1998	BS chemistry/biology	BS minor in education MEd education Science teacher, K-9 Middle school principal for 15 years	
UC Berkeley	C.K., 1998	BA mathematics	Community college teacher of computer science	MBA Work experience in mathematical programming
	M.L., 1998	BS nutrition MPH nutrition		Some experience in survey research Contract monitor for large nutrition education evaluation

INSTITUTION	Fellow, year of ETP entry	Science/math background	Education background	Other relevant background
[UC Berkeley, con′t]	M.T., 1998	Experience and expertise in technology		BA geography Work experience in measurement and student assessment
	A.B., 1999			BA psychology MA experimental psychology Prior BEAR evaluation work
	C.S., 1999	BS microbiology	MA education Science teacher	Some experience with classroom testing
	D.S., 1999	BS mathematics MS statistics	Math teacher, middle, secondary, community college, and university levels	
University of Minnesota	D.B., 1997	BS chemistry	MAT science education Science teacher, 9-12	Work in private industry as a chemist
	T.F., 1997	BS mathematics education	BS mathematics education MA educational administration Math teacher, high school	Participated in summer evaluation institute at Western Michigan Univ
	B.J., 1997	BS biology education	BS biology education MA education Biology teacher, high school	Experience in curriculum evaluation
	L.C., 1998	BS math	BS physical education MA education (C&I) Math and physical education teacher, junior and high school and college levels	

INSTITUTION	Fellow, year of ETP entry	Science/math background	Education background	Other relevant background
[Univ of Minnesota, con′t]	A.M., 1998			BA psychology and sociology MA educational psychology Graduate coursework in statistics, evaluation, and measurement
	B.L., 2000	BS biology MS ecology (in process)	Teaching certificate Biology and chemistry teacher, high school	Peace Corps volunteer
Utah State University	C.E., 1997	BS computer science, minor in chemistry	MS instructional technology University instructor	Worked in private industry as engineering technician, quality control chemist, systems programmer, computer network administrator, and other jobs
	L.M., 1997	BS biology, minor in chemistry	MEd curriculum and instruction Science teacher, high school Math/science curriculum coordinator, high school	
	N.D., 1997	BS mathematics, computer science, statistics	Teaching certificate Teacher, 6-12, math and computer science	Active in professional math education associations
	T.B., 1998	BS chemistry, minor in biology MS civil and environmental engineering		BS psychology MS family and child development Experience with MIS and research at Center for Persons with Disabilities

INSTITUTION	Fellow, year of ETP entry	Science/math background	Education background	Other relevant background
[Utah State Univ, con′t]	R.G., 1998	BS biology	BS secondary education Science teacher, high school	
	K.W., 1998	BS geology and hydrogeology	MS instructional technology Science teacher, high school	Research experience at Idaho National Engineering Lab
	M.N., 1999	BS math/statistics MS applied statistics	Developed and taught university courses in use of statistical software	
Program withdrawals	T.D., 1998 — *Minnesota*		BA elementary education MA elementary education with science emphasis Science teacher, fifth grade, middle school	Assistant director of education outreach for a NASA Mars voyage program
	N.P., 1997 — *Utah*	BS mathematics, minor in biology	MEd curriculum and instruction Math/science teacher, junior and senior high school, university levels	

Table 6
Jobs and Job Domains Desired by ETP Fellows*

What do you hope to be or anticipate doing after you complete the ETP?	**Do you hope or expect to be actively involved in science/ mathematics education programs?**	**Do you hope or expect to be actively involved in evaluation?**
Act as an internal evaluator for community-based programs engaged in environmental action	Yes, this will be a major aspect of my work.	Yes, this will be a core element of my work.
Consult with schools on curriculum and on program selection and implementation, evaluate curricula in both science and math	Yes, that is my goal. I am comfortable working in math/science, although I currently work in all subject areas.	I would like to do most of my work in schools, yet also continue the two evaluation projects I am currently involved with.
Continue to work on issues of equity and multiculturalism in science/math programs and systemic change initiatives	Yes, I hope to include pre-service teacher education programs in my work.	I hope so, including evaluations of teacher education programs at state teachers colleges.
Work in an evaluation and research unit within a center devoted to research on science and science education	Yes	I hope so. I hope for a full time evaluation job with a science research center.
Evaluate educational projects for a larger organization.	Yes, I would choose to work in science education as that is my strength.	I hope so.
Evaluate community college educational programs, probably in distance education	Yes, I will be working with mathematics and computer science education at the high school and community college levels.	Yes, I will be evaluating classroom innovation as well as online education
Be an evaluator for the XY State Department of Education	Yes, I will be actively involved in evaluating nutrition and nutrition education programs for the State of XY.	Yes, up to 50% of my full-time job, maybe more.
Run statistical analysis for evaluation/research projects	I am still very interested in science and math reform efforts.	I anticipate being involved in evaluations as a behind-the-scenes person doing statistical analysis, rather than being a PI at this point.
Work as a postdoc	Yes, the postdoc is in mathematics education.	Yes, I have been contacted to do two evaluations so far.

What do you hope to be or anticipate doing after you complete the ETP?	Do you hope or expect to be actively involved in science/ mathematics education programs?	Do you hope or expect to be actively involved in evaluation?
Finish my degree and then go fishing	Yes, but I would like my involvement in these programs to be a secondary part of my professional life, not my primary job.	Yes, but I would like my involvement in evaluation to be a secondary part of my professional life, not my primary job.
Start my new job as an assistant professor of science education	Yes, I expect extensive involvement.	Yes, I expect extensive involvement.
Teach at a community college and eventually be a PI of a science education project at my community college	Yes, I hope to participate in and eventually run a science education program.	I would build evaluation into my program, but someone else would do the evaluation.
Work for a not-for-profit group in educational evaluation/research	If these programs are part of the organization for which I will work.	Yes, definitely.
Consult	Yes, because my focus is educational research and because science/mathematics education has become a national priority, I expect I will be active in evaluating these kinds of programs.	Yes, evaluation work in my career.
Be a university faculty member	I hope to have the opportunity to use my evaluation skills in evaluating educational programs in math/science.	I hope to have regular opportunity to use my evaluation training in a variety of capacities.
	Yes, I expect to be very actively involved.	Yes, I expect to be very actively involved.
Evaluate science programs and classroom teachers in the classroom practicing their craft. I also want to be published, to be a scholar.	Yes	Yes, by asking evaluative questions of teachers and other school people, and by doing a few contracts on the side.
Be a faculty member in math education. I don′t feel confident doing evaluation.	Yes, I will teach math ed	I would like to do evaluations also, say, as a consultant.

* These responses are extracted, verbatim or as close paraphrases, from questionnaire and interview data.