Corridors for Collaboration

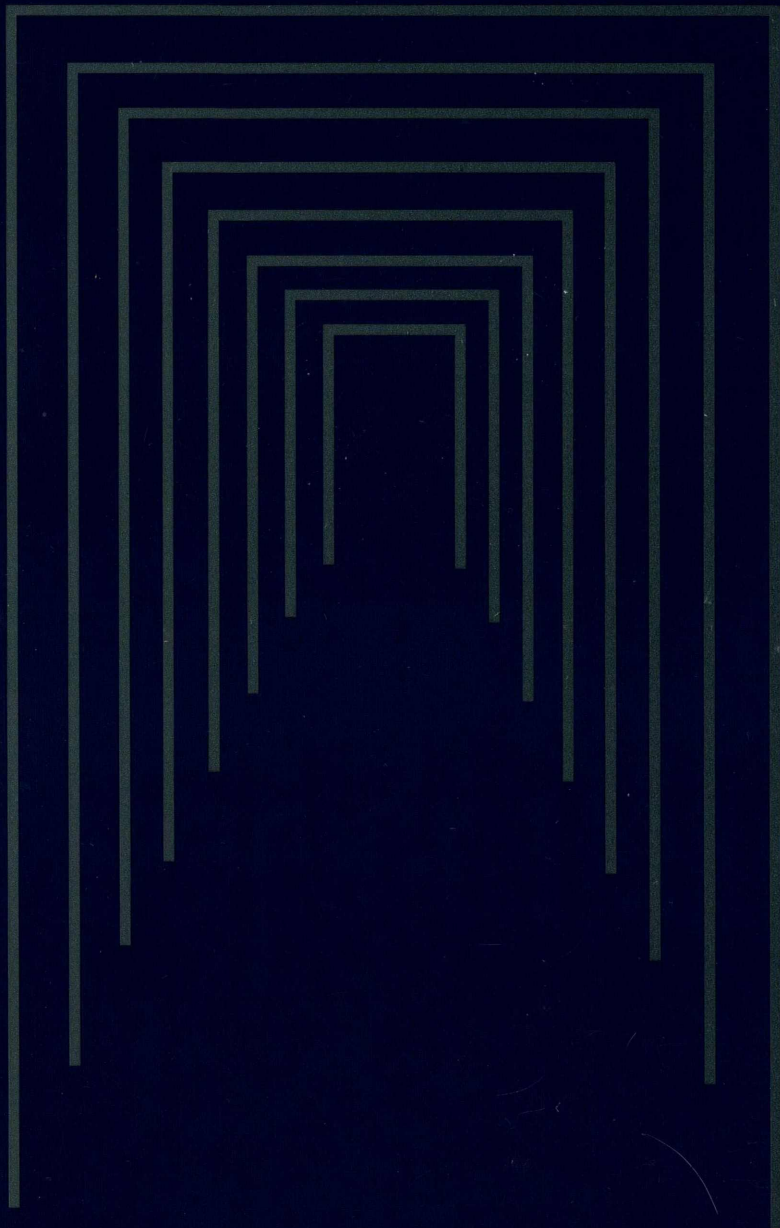# Corridors for Collaboration

Collaboratives for Excellence in Teacher Preparation (CETP)
The National Science Foundation

Evaluators' Meeting

Arlington, Virginia
February 4-6, 1998

Prepared by

Carl Frantz
Director of Research and Sponsored Programs
University of Southwestern Louisiana

Frances Lawrenz
Professor of Science Program Evaluation
University of Minnesota

Jennifer Kushner and Susan B. Millar
Learning through Evaluation, Adaptation and Dissemination (LEAD) Center
University of Wisconsin-Madison

Table of Contents

Introduction to the Proceedings and
Outline of Directorate and Division Priorities

Norman Fortenberry
Division Director
Division of Undergraduate Education
Directorate for Education and Human Resources
National Science Foundation

It was a pleasure to read the proceedings of the first joint meeting of the evaluators of the many NSF Collaboratives for Excellence in Teacher Preparation (CETP) projects in the nation. CETP is an important component of the Division of Undergraduate Education (DUE) portfolio and evaluation is becoming an integral component of all our projects and programs.

DUE is part of the Directorate for Education and Human Resources (EHR) of the National Science Foundation (NSF). The Directorate is responsible for providing national leadership and support for NSF's efforts to improve the quality of science, mathematics, engineering and technology (SMET) education, kindergarten through graduate school. DUE serves as the focal point of NSF's agency-wide effort to foster improvement in undergraduate SMET education for all students (SMET majors, prospective pre-kindergarten through grade twelve teachers, students preparing for the technical workforce and all students as citizens in an increasingly technological society) in all types of institutions (two-year and four-year colleges, as well as comprehensive, doctoral, and research intensive universities). In exercising their responsibilities, the Directorate and the Division have established the SMET education of the nation's teachers as one of their highest priorities. The ultimate goal is to produce teachers who are knowledgeable in the content areas and in the practice of teaching, creative and enthusiastic, and dedicated to life-long learning.

Beginning in FY1999 there will be three programs in DUE, each of which will accept projects concentrating on teacher preparation: CETP which concentrates on comprehensive reform of the entire continuum of a teacher preparation program; Advanced Technological Education which concentrates on two-year institutions and their role in producing the nation's technological workforce; and the Course Curriculum and Laboratory Improvement Program which includes efforts to support development of new materials and practices, adaptation of proven innovations, and national efforts to improve the quality of SMET instruction by current and future faculty.

This is an exciting time for those of us concerned with improving the SMET preparation of the Nation's teachers. However, the opportunity also creates a responsibility to not only carry out that change in an exemplary manner, but also to determine if the changes instituted by each project and the overall program have been successful in achieving the overarching goals; well prepared teachers and a well defined teacher preparation program that has become the acknowledged responsibility of the institutions involved. This meeting was conducted and designed under the able leadership of evaluators who have been charged with project evaluation from the very

beginning of the program, Carl Frantz of the Louisiana Collaborative for Excellence in Teacher preparation and Frances Lawrenz, a consultant to numerous sites, and with the input of evaluators who have been working with other large systemic projects focused on undergraduate education, Susan Millar of the Chemistry Initiatives and Theresa Smith of the Alliance for Minority Participation. The purpose of the meeting is to enable cross fertilization of ideas and exchange of instruments to promote more uniform measures across the collaboratives and to inform the attendees of changes in NSF programs. These proceedings are important not only because they offer an opportunity to exchange information concerning evaluation strategies and instruments, but also because they help to further refine various initiatives at NSF; chief among them, the needs of the Government Performance and Results Act. This is a mandate from Congress for each federally funded agency to define its goals in easily measurable terms. For the collaboratives this translates into the need to document outcomes in three areas: the teachers produced; the programs produced; and, the nature of the collaborative efforts and their effect on institutional attitudes and practices.

TEACHERS PRODUCED

What evidence exists that the changes instituted have indeed resulted in a substantial increase in the number of students who know more, who are more comfortable with teaching mathematics and the sciences using the mathematics and science standards as a guide, and who are comfortable with employing the new technologies available?

DEFINITION OF PROGRAM

We have given each site considerable latitude to define its own program but that does not mean there is no need to define the new program in terms of:

1. What at the end of five years will be the changes in the way teachers are produced within the institutions participating (e.g. change in course requirements, increased coordination of methods and content courses; earlier or increased field experiences?).

2. What experiences are the minimum definition for a student to be regarded as a product of a specific CETP program?

3. What are the minimum requirements for certification at the institutions involved? Do these exceed those of the state or did they help frame current state requirements?

INSTITUTIONAL OUTCOMES

What documentation exists for changes occurring in institutional practice and attitudes as regards teacher preparation? I commend you for the hard work you have done to prepare for this meeting and thank Frances Lawrenz, Susan Millar and the Louisiana Collaborative, in particular Carl Franz, for the effort they expended to coordinate this event and the production of the proceedings. These proceedings will serve as an

excellent basis for developing evaluation strategies useful to 1) faculty revising courses, 2) departments and institutions changing their teacher preparation programs and their campus culture to increasingly recognize the importance and needs of students preparing to be teachers, and 3) national leaders within public and private sectors interested in the results of support for comprehensive and systemic reform.

Evaluation and GPRA Within EHR

Conrad Katzenmeyer
Senior Program Director
Division of Research, Evaluation and Communication

Evaluation has had a prominent role in the Education and Human Resources Directorate for some time, but particularly since 1992 when Congress mandated the Directorate to initiate evaluations of all of its programs over a five-year period. This has occurred, and EHR staff awareness of evaluation needs has grown markedly.

The same has been true for Project PIs (Principal Investigators) and others in the field. We in the Evaluation Program have noted a marked increase in the knowledge that PIs bring to evaluation workshops we conduct. A few years back, discussion of formative vs. summative evaluation was a major topic of concern. Now participants start from a position well beyond this level of knowledge. Further, we have heard from others in the field that NSF has the reputation of requiring strong evaluations with good outcome measures. This is a big step forward from how NSF's requirements for evaluation were once viewed.

It is very likely that the Government Performance and Results Act (GPRA) will accelerate the emphasis on evaluation. Passed in 1993, GPRA requires that Federal agencies determine measurable goals for all of their program activities. Agencies must then measure their performance against these goals and report progress to Congress as part of the annual budget submission. The first of these reports is due in FY 2000 and NSF is on track to submit this report as required.

One outcome of the GPRA process will be greater emphasis on regular collection of information. Since GPRA is to be on a yearly cycle matched to budget submissions, it is reasonable to expect that new outcome information will be required each year. While there may be exceptions, most EHR programs can expect a requirement of updated information each year. GPRA will also place an emphasis on collection of quantitative data. While not limited to quantitative data, the accountability aspect of GPRA will probably place a premium on these data. Perhaps most importantly, GPRA will require that management of programs be aligned with outcome data. How this will be done is yet to be determined as the GPRA accountability system is established, but certainly EHR management has already placed greater emphasis on quantitative outcomes when they look at program results, as those who have been through a Program Effectiveness Review can attest.

GPRA will also force EHR to be clear about the kinds of data that are to be collected, and for what purposes. Distinctions such as the following need to be made.

Project evaluations. Project evaluators will continue to be concerned primarily with devising and collecting data on specific project objectives and outcomes. The major audiences will be local and the field of SMET education, although NSF is likely to express greater interest in these results than it has in the past as part of its accountability responsibilities.

Program monitoring. The intent of program monitoring is to provide Program Officers with information they need to manage the Program. Data will likely be collected from all projects within a program around a common set of indicators. Emphasis will be on outcomes collected by computer, but process and qualitative data may be included, as well.

Program evaluations. Program evaluations will emphasize summative judgments of the overall effectiveness of EHR programs. The data required would relate to the program as a whole, and may be identical to some of the indicators collected for monitoring. But other program evaluation data might be drawn from a sample of projects, it might be gathered using data techniques (e.g., observations and standardized tests) that are not being used by particular projects, and may include variables (e.g., impact on the general field of science) that are not being addressed by any of the individual projects in the program. Unlike program monitoring data, program evaluations are likely to include comparison group data.

GPRA. The existence of the GPRA legislation adds another requirement on collection of data. GPRA reporting is almost certain to be highly aggregated and thus the data requirements are likely to be less than in other categories of data collection. Hopefully, many of the GPRA data needs can be filled by data from these other collections, but not necessarily so. GPRA will also impose its own timetable for collection.

For the sake of efficiency, one of our goals must be to keep data collection as lean as possible, and to use data collections for multiple purposes whenever possible. Another goal is to maximize communication among the interested parties – project and program staffs, external evaluators, and evaluation staff in EHR – so that we take advantage of our accumulated knowledge in designing good data collections.

One point is certain. Given the resources available, data collected by contractors can only meet a minor portion of overall NSF data needs. Whatever the purpose, projects will often be called upon to collect data beyond what was previously expected because the Foundation has no field locations or staff to carry out this function. It is essential that we all recognize the requirements that will be placed on projects, requirements that did not exist just a few years ago, and that adequate resources be made available for meeting these new requirements.

Executive Summary

This report describes the proceedings of the February 1998 meeting of National Science Foundation (NSF) Collaboratives for Excellence in Teacher Preparation (CETP) project evaluators, program managers and external NSF funded evaluators. The meeting purpose was to facilitate the goals of CETP which are to "improve significantly the science, mathematics, and technology preparation of future K-12 teachers and their effectiveness as educators in these areas" (*CETP Program Announcement*, NSF 97-29) through increasing the effectiveness and consistency of CETP evaluations. It was a very productive meeting with a positive orientation toward the benefits of collaboration among the various players in the CETP program and projects. Through both large and small group discussions, meeting participants identified several avenues or corridors for collaboration and offered many suggestions for collaboration within these corridors. Most important were corridors among NSF program evaluation and the individual project evaluations, corridors among the various players within the CETP projects and corridors among the CETP projects themselves.

The project and program evaluators believed this meeting provided an excellent opportunity for both to learn more about, as well as inform, each other. Project evaluators want to know how their projects fit into overall program effectiveness and find ways to better align their goals with overall program evaluation goals. Additionally, project evaluators would like to have closer working relationships with the program evaluators and would like to consider themselves a valuable resource for program evaluators in terms of providing assistance in collecting data, sharing instruments, or assisting in program evaluation planning. Participants agreed it is important that program and project evaluators share the different types of data each are collecting in order to maximize the use of the limited evaluation resources.

Creating a positive, informed relationship between the PI and evaluator(s) is extremely important. The PI must understand and agree to the evaluation strategy and responses to key evaluation issues at the outset. Also important for evaluators is the process of developing the trust and confidence of involved professors and other key participants on each campus.

The CETP Evaluators' Meeting provided many potential opportunities for collaboration across projects. Besides creating additional lines of communication generally, evaluators identified various potentially valuable joint evaluation initiatives. Even though much diversity occurs across CETP projects, the similarities in purpose and activities do provide real opportunities for 1) sharing of information, insights, and experience about what works and does not work and 2) joint activities.

This report highlights several opportunities and issues discussed at the meeting.

Introduction

The National Science Foundation sponsored a February 4-6, 1998 meeting for project evaluators from the various Collaboratives for Excellence in Teacher Preparation (CETP) projects across the nation. Dr. Terry Woodin, CETP Program Director, Dr. Conrad Katzenmeyer, Research, Evaluation and Communication (REC) Senior Program Director, and other NSF officials, external evaluators, and other interested individuals (see Appendix B for a list of meeting participants) joined them. The CETP program is an education reform initiative developed by the NSF's Division of Undergraduate Education (DUE). CETP's goals are to, "improve significantly the science, mathematics, and technology preparation of future K-12 teachers and their effectiveness as educators in these areas" (*CETP Program Announcement*, NSF 97-29).

Purpose

The meeting's intent was to help identify how evaluation could facilitate CETP program goal attainment. The meeting's specific purposes were to 1) increase the effectiveness and consistency of CETP formative and summative evaluations, and 2) facilitate collaboration and networking among CETP and related evaluators. In addition, the organizers designed the meeting to inform CETP project evaluators about NSF expectations, especially those regarding summative evaluation. The meeting proceeded with a series of large and small group sessions where participants discussed evaluation issues, challenges and solutions. See complete agenda in Appendix C.

Outcomes

The meeting provided the opportunity for valuable discussions among the CETP project evaluators, NSF evaluators and NSF program managers. The possible ways in which collaboration could increase the impact of evaluation were emphasized at the meeting. Participants discussed the major types (or corridors) of collaboration:

- corridors between NSF and the CETP projects;
- corridors within CETP projects; and
- corridors between the CETP projects.

Additionally, participants identified ways to enhance existing collaboration activities and to develop new mechanisms for collaboration.

The NSF/CETP Projects Corridor

Government Performance and Results Act (GPRA)

Dr. Norman Fortenberry, Dr. Conrad Katzenmeyer and other NSF meeting participants reported heightened interest in evaluation because of the Government Performance and Results Act (GPRA). GPRA mandates that measurable outcomes be used to manage all federally funded programs. The outcomes of GPRA-mandated summative evaluation, thus, will influence how funding is allocated for programming in the future. GPRA-designated accountability indicators include evidence for discoveries, connections, a well-trained workforce of scientists and engineers, and improved achievement in mathematics and science. Consistent with these new accountability guidelines, NSF will place more emphasis on databases, project and program monitoring, impact studies, project and program reviews, and communication. NSF will ask the projects to help respond to the new GPRA demands because the ultimate sources of data are people involved in the projects.

This meeting provided the opportunity for CETP project evaluators to suggest and consider how their current and future evaluation activities could help meet NSF needs. Consistent with GPRA requirements, evaluation at the program level has shifted to a more summative approach. Project evaluators, thus, are interested in ways to balance program expectations for summative data with the largely formative evaluation approaches used in the projects. Participants agreed that NSF program evaluation information needs should be communicated as soon as possible to allow project evaluators the time to modify project evaluation plans to address those needs.

NSF Division of Research, Evaluation and Communication (REC)

REC is responsible for obtaining evaluative data on the entire Directorate for Education and Human Resources programs and has two contracts with SRI International for evaluation of the CETP program. One contract was let before GPRA and the other is just beginning.

As previously noted, a central CETP goal is to significantly improve the preparation of future K-12 mathematics, science and technology teachers. A primary strategy NSF suggests for achieving CETP goals is to create linkages between institutions of higher education and public school systems, as well as between the disciplines of mathematics and science. In keeping with CETP goals and strategies, NSF and Stanford Research International (SRI) have identified indicators of CETP program success for their summative program evaluation. These include changes in the number and quality of the teachers produced and evidence of institutionalization, as measured by the nature of the relationships among partners within a Collaborative or among departments within or across institutions, institutional support for teachers, and changes in the content and instruction of courses.

The original program evaluation effort by SRI included one- or two-day annual site visits to the various CETP projects. The project evaluators made three main suggestions for improvements in this process to enhance the usefulness of SRI site visits.

- First, because some project evaluators were unclear about what the site visits are intended to accomplish, participants suggested that the objectives (and indicators being used to measure success) be clarified and communicated to the local projects.

- Second, they suggested that before leaving the site, teams should share insights they gained from the visit informally with the Project PIs or project evaluators in the spirit of formative evaluation and clarification of goals.

- Finally, the project evaluators suggested that the visits be scheduled to optimize information gathering and minimize intrusion. Teams could visit when other external review teams were visiting or in concert with existing events so evaluators could "see things in action". In this regard, participants suggested that it would be ideal if visiting teams were comprised of pedagogy experts and classroom practitioners with expertise in both mathematics and science. These experts should be familiar with the project context before arriving at the site. Also, project evaluators felt it would be useful if more time and/or additional sites could be included.

For their part, SRI program evaluators conveyed that they understood the project evaluators' recommendations about the site visits. At the same time, they explained that SRI might not have the resources to respond adequately to all of these suggestions. SRI evaluators feel stretched to visit all sites within a collaborative even once during the five-year grant period.

Meeting participants agreed that common indicators of success are useful in assessing the CETP program overall. At the same time, the group agreed that since different models were funded in the array of CETP projects, the program might implicitly be accepting multiple definitions of success or at least multiple ways of achieving it. In particular, the project evaluators discussed how differences among the Collaboratives' strategies for achieving program goals are accompanied by differences in the types of outcomes they are seeking to measure. The differences among collaboratives make it difficult for all projects to provide data on a common set of identifiers and for SRI to have overall indicators. Along this same line, a number of the project evaluators suggested that using comparison groups across projects in SRI's evaluation would be difficult because the project contexts are so diverse.

In order to overcome this difficulty, both program (SRI) and project evaluators agreed that the process of defining program success and designing an appropriate program evaluation should be guided by questions that recognize the unique characteristics of the individual projects and allow for summative statements about program impact. Questions such as the following were suggested: What are we supporting? How are we supporting it? What are we learning? What are we doing with what we learn?

Project evaluators felt it would be useful for SRI to send them interim reports or a summary of the final report on the program. While summative program evaluation has traditionally been for internal NSF use, project evaluators believe that the program evaluation reports would help them understand how their efforts and objectives are fitting into the bigger "program" picture. Moreover, access to these findings would help project evaluators and their partners identify ways in which they are duplicating the program evaluators' efforts, as well as ways in which data and resources could better be shared. It would also help with the communication and dissemination efforts required by GPRA.

NSF Division of Undergraduate Education (DUE)

To conduct its own program evaluation of CETP projects, DUE is involved with three major evaluation efforts: monitoring, Quantum Research Corporation's (QRC) databases, and the National Visiting Committees. These offer several opportunities for collaboration between DUE and the projects. The monitoring process involves collaborative communication with the NSF program officers in DUE. Participants encouraged DUE to spend more time in this sort of ongoing information exchange because it would facilitate timely formative project improvements rather than focus on the more summative, labor and time intensive annual reports or reverse site visits. The planned faculty on-line surveys and telephone surveys with "lead" players at CETP institutions are a step in this direction.

Project evaluators thought DUE planned analysis of information from the QRC databases would be informative. Since QRC began its data collection, NSF and QRC have been responsive to CETP projects' requests to work more collaboratively. This has resulted in a collection system that is more user-friendly and more worthwhile. However, at least some project evaluators felt they would benefit from more specific descriptions of the data requested in the QRC instrument to complement the clearly presented procedural instructions.

Most importantly the projects would appreciate receiving data supplied by their own projects and, to the extent reasonable, aggregated data from other CETP projects in spreadsheet format. This would allow them more flexibility in analyzing QRC-collected data from their own project and would provide more insight concerning how their project compares with the other CETP projects.

The National Visiting Committees are another device for providing both formative and summative information. These committees have served useful roles in the projects and recently DUE clarified their expected role. CETP project evaluators indicated, however, that it is difficult for the committees to function well in both formative (helpful) and summative (judgmental) manners. The committees also bring another layer of complexity to projects that are already attempting to meet a variety of evaluation information needs.

In sum, the project and program evaluators believed this meeting provided an excellent opportunity for both to learn more about, as well as inform, each other. Project evaluators want to know how their projects fit into overall program effectiveness and find ways to better align their goals with overall program evaluation goals. Additionally, project evaluators would like to have closer working relationships with the program evaluators and consider themselves a valuable resource for program evaluators in terms of providing assistance in collecting data, sharing instruments, or assisting in program evaluation planning. All participants agreed it is important that program and project evaluators share the different types of data each are collecting in order to maximize the use of the limited evaluation resources.

This meeting was a major step in establishing common objectives and indicators since all participants better understand each other's needs. The project evaluators recommended that the progress they had made during the meeting be incorporated into the system of communication between program and projects. Improved communication will help the projects stay abreast of the changes at the program level and enable their evaluation efforts to better inform program evaluation.

"Within CETP Project" Corridors

Expectations for CETP project evaluation and evaluators are diverse because each CETP project is unique and complex. Exploring this diversity revealed many opportunities for collaboration within each CETP. Optimizing collaboration at this level is critical for support of the evaluation efforts and effective use of evaluation data. The major opportunities occur among the PI's and the evaluators, the project participants and the evaluators and among the project participants themselves.

Project PI's and Evaluators

Clear understandings between the Project PI's and the evaluators are necessary for effective evaluation. Evaluators from the first two CETP cohorts underlined the importance of evaluators spending time early in the project with the PI for two purposes—developing a common vision and helping the PI understand the value of evaluation. They emphasized that this time investment would minimize potential difficulty later and set the stage for an effective working relationship. Evaluators were encouraged to spend time investigating the barriers and resolving the issues that create tension. Collaboration between evaluators and PI's appeared to be strongest in projects in which it was considered the role of both to foster collaboration within the project.

In this same vein, some meeting participants felt that "housekeeping" issues regarding evaluation data and products should be addressed proactively. The main housekeeping issues to consider involve ownership of the data, audience for the evaluation reports, and report formats. Prior to conducting the evaluation, the project evaluator and PI should be in agreement about the purpose of the evaluation, the roles the evaluator will be expected to play, definitions of project success, and the limits of the measurements.

The process of negotiating questions about roles provides an opportunity for collaboration. The questions that could shape these kinds of discussions include:

- Should an evaluator be involved in setting project/assessment objectives?
- Is it the role of the evaluator to help the PI continuously articulate project goals?
- Should an evaluator press for more resources for evaluation?
- Can an evaluator be "too internal" and thus lose third party perspective?
- Can the evaluator report directly to the PI?

A number of project evaluators noted the importance of a well-defined and smooth working relationship with the PI. They stressed the need for the PI to provide adequate financial and political support for evaluation. One difficulty for evaluators in fostering collaboration is sharing negative information. Since most evaluators see their role in terms of facilitating project improvement, they were uncomfortable being perceived this way. Project evaluators who had been in this position shared with the group a number of strategies they used to communicate "critical" feedback. The most common approaches involved first sharing the information with the PI, using non-

judgmental descriptions, and stating assessment findings positively to the broader audience.

Most evaluators would like more opportunities for PI's and evaluators across the projects to get acquainted. They suggested that PI-evaluator teams from each project visit other sites to become familiar with other PI's and evaluators.

Evaluators and Project Participants

The evaluators suggested spending substantial time developing and maintaining good working relationships with the faculty members, administrators, other project members and institutions involved in their projects. To help accomplish this, evaluators need to demonstrate and explain the value that their efforts add to a project. Once developed, these relationships provide the trust necessary for evaluators to obtain necessary help and cooperation from faculty members and others. Obtaining data from the various participants in the Collaboratives can be difficult because those involved lead very busy lives with other priorities and issues. By developing strong relationships early, evaluators can facilitate data collection and help the project.

A number of the evaluators considered the first three project years to be particularly important in developing relationships and establishing trust because the primary focus then is on formative evaluation for project improvement, not project justification. One evaluator felt that she built trust by making clear to others that she had their best interests at heart. This requires being open to differences and being sensitive to the issues important to different players.

Working closely with participating institutions at the outset often facilitates the use of existing institutional data, such as student demographic data, grades, and graduation rates, that are useful to project evaluators. These data can be used to produce longitudinal databases, and are particularly useful for "big picture" understandings of project success. Evaluators can use existing institutional data for comparisons between reform courses and traditional courses and also to capture the curricular paths and academic outcomes of an entire student cohort.

Some evaluators felt that to foster collaborative relationships, faculty, PI's, evaluators and teachers need early opportunities to meet, become familiar with each other's areas of expertise, find common sets of goals and a common language, and get to know each other as people outside of their professional roles. Other helpful hints included: bartering work with other project evaluators, involving students in the evaluation process as helpers, finding common meeting places, negotiating common definitions, and being aware of any words or phrases that would be perceived as alienating.

Evaluators discussed instances where some participants were not supportive of the evaluation activity. Traditional faculty or others who are unconvinced of the value of reform may not understand and support the project, let alone the evaluation effort. Some may even feel threatened by assessments of student outcomes. These persons may express their resistance by not participating in evaluation activities that are aligned with the project reform initiatives. Evaluators can play an important role in building support

by discussing areas of resistance and by playing the role of a mediator who listens carefully to these faculty members.

Evaluators can also help to build collaboration and support for the project by sharing their expertise with project members. Two suggestions in particular are 1) for the evaluator to help inform faculty members about new methods of assessment and 2) for the evaluator to help faculty members design "action research".

Among Project Members

One of the major CETP program objectives is to promote, strengthen and solidify lasting linkages among partners in the CETP projects. This objective is essential to create an entire system that brings more and better-educated teachers into the mathematics and science teaching force. Project evaluators discussed the need to go beyond reporting numerical indicators of change to identifying the structural and contextual features of these Collaboratives that enable them to accomplish their objectives. Evaluation efforts useful to study collaboration were mostly qualitative and focused on describing the relationships within an institution among project partners. Methods used included observations, interviews, focus groups and surveys. Studies of the collaboration process often help to improve it, because people are encouraged to reflect on the process and what it means.

Evaluators in the first two CETP cohorts presented a number of particular strategies that they used to foster collaboration. These are briefly outlined below.

> Maryland Collaborative for Teacher Preparation: The Maryland PIs worked at collaboration by providing opportunities for partners to get to know each other. They did this by hosting meetings for everyone. The PIs and evaluators recognized that it takes a long time to develop trust and began the process by trying to develop rapport and common interests. The PIs and evaluators planned meeting agendas together. Both evaluation and research teams were written into the grant as distinct entities. Role definition within both these groups required continuous clarification. The Collaborative also used distance technology (video conferencing) to bring people from many institutions together.

> Louisiana Collaborative for Excellence in Preparation of Teachers: The initial evaluation efforts were formative in nature—emphasizing "lessons learned" to improve campus renewal efforts rather than imposing external student outcome measures. Site visits to each of fourteen campuses were held in which interviews, focus group discussions, and classroom observations were conducted. Project evaluators of LaCEPT provided support, encouragement and suggestions to faculty to conduct their own outcome-related assessment in their classrooms.

> Montana STEP Project: STEP evaluators made site visits within the Collaborative to learn about the interests other partners had for the evaluation. These conversations were very important for building relationships and helped to

understand why people were or were not participating. Evaluators created a matrix that illustrated the relationships between institutions and people within the Collaborative and found that to be helpful in sharing the 'big picture" with others.

<u>Rocky Mountain Teacher Excellence Collaborative:</u> Evaluators are internal, which helped address some of the trust issues. The evaluators fostered collaboration among people in like disciplines but separate institutions by providing the opportunity for group meetings.

<u>Philadelphia Collaborative for Excellence in Teacher Preparation:</u> The evaluators from this CETP emphasized that they created a plan for evaluation which took into account the short windows of opportunity for various players to communicate with each other. Their PIs acted as "marketers" of the program, which took pressure off of evaluators. The evaluators intentionally used the process of evaluation to bring people together.

<u>New York Collaborative for Excellence in Teacher Preparation:</u> New York evaluators fostered collaboration between faculty of like-disciplines located at different sites within the Collaborative by creating a round-robin mentoring program.

<u>Virginia Urban Center Teacher Preparation Collaborative:</u> This Collaborative hosted summer colloquium that used evaluation to train new teachers.

Summary of the "Within CETP Project" Corridors

Creating a positive, informed relationship between the PI and evaluator(s) is extremely important. The PI must understand and agree to the evaluation strategy and responses to key evaluation issues at the outset. Also important for evaluators is the process of developing the trust and confidence of involved professors and other key participants on each campus.

The "Across CETP Project" Corridors

Methodology

Much of the CETP evaluators' meeting discussion focused on methodology. The wealth of evaluation experience available at the meeting was evident and the discussions emphasized how much the evaluators could learn from and help each other.

The project evaluators identified various factors that shape their choices about methodology including project objectives, available resources, personal expertise, personal preferences, audience, and timetable. They gave special attention to CETP project objectives, emphasizing that each project has multiple dimensions that need to be evaluated. For example, most CETP initiatives seek to change course content, methods of instruction, attitudes about teaching, attitudes about change, and university support for improving the teaching of mathematics and science. Moreover, the evaluators noted that they must employ different evaluation designs in order to accomplish both the formative and summative objectives of their work.

CETP project evaluators emphasized the importance of employing multiple methods to evaluate their projects. They felt this to be especially important due to their projects' complex natures and to meet the needs of the many audiences concerned about evaluation results. They stressed that multiple methods allow evaluators to collect different types of data about all aspects of project components from qualitative, rich descriptions of phenomena to quantitative statistical analyses. To optimize efficiency and avoid duplication of effort, PI's and evaluators should be clear about what data will be collected in common across all CETP projects and whether there are any common instruments for collecting these data.

All the projects are operating in environments where there are factors in place that are beyond the control of the project but significantly affect attainment of project goals. Examples of such factors include state standards for achievement in K-12, state- or district-mandated teacher salary levels, and state-mandated teacher preparation activities. Furthermore each Collaborative is comprised of numerous institutions, each with its own mission, resources and culture. The project evaluators emphasized that conducting evaluation at multiple sites not only requires significant logistical effort, but also sophisticated understanding of the specific context of each institution. A special challenge is to identify indicators of change that are appropriate to all institutions.

Likewise, evaluators may find it difficult to establish common criteria for baseline data since each institution may have different stakeholder expectations (school boards, parents, faculty) and different student bodies, faculty cultures and reward systems. It is essential they take these factors into account in designing evaluation plans and in reporting evaluation findings because they may significantly impact project outcomes.

Among the several ideas about different methodologies that were presented were the following.

- However difficult it may be to collect common baseline data, such data are extremely important nevertheless. Even when "control" or "comparison" groups are hard to identify, CETP sites can compare faculty and student outcomes and other factors after CETP reform interventions ("treatments") with what existed before those interventions. The Philadelphia Consortium, for instance, had the foresight to begin collecting baseline data about faculty and student attitudes, retention rates, grades, and other relevant outcomes even before being officially awarded the grant.

- A significant factor in achieving change, whether within a sub-system or system-wide, is the degree to which key players are invested in the reform process. This kind of investment tends to be difficult to measure. Methods of assessing commitment to change used by evaluators present at the meeting include assessing faculty attitudes about the reform process specifically asking faculty and administers to rate traditional objectives in comparison with reformed objectives.

- In a number of the projects, institutions or school districts are involved in multiple reform initiatives. At these sites, multiple interventions are occurring simultaneously. This can create complications for evaluation in that there may be multiple effects of overlapping programs, making it difficult to stipulate which effects were caused by which interventions. The evaluators reported that they are doing their best to address this difficulty, using multiple evaluation methods to both describe and measure changes so that the confounding among the different interventions can be identified. Dr. Terry Woodin assured them that she understands these difficulties and encouraged them to acknowledge confounding effects by simply stating the limitations of their measurements. The *Western Massachusetts Collaborative (STEMTEC) (Appendix C)* offered some advice on how to attribute change to an intervention when there is only an incomplete definition of what the intervention is, and on how to look at interventions on a continuum while considering contextual issues.

- One of the most complex methodological issues addressed at the meeting was how to measure longitudinal change, including long-term impact on student learning and retention of reformed teachers in the teaching force. Presently project evaluations provide only intermediate indicators (predictors) of long range effectiveness. These efforts need to be used in conjunction with other approaches when measuring longitudinal change. In this regard, some evaluators felt that the impact of the reformed courses and other elements of the CETP programs may not be noticeable in the time period allotted for evaluation. In particular, changes in attitudes and understandings about teaching careers may not be apparent for years.

To understand the impact of the project overall, many evaluators felt that it is important to interview graduating pre-service teachers. They also felt it important to talk with in-service teachers who have participated in a CETP project to learn how their perceptions change over time. While interviews could be used to gain this information, another interesting way to accomplish an evaluation of this type is to conduct ethnography in order to distinguish the effects of the CETP program from the many other major factors affecting these teachers (such as school culture/context).

Course Reform

Determining what constitutes a reformed course is an issue of critical importance to all of the CETP projects. Project evaluators were asked to share their views of the features the CETP projects' leaders expect should characterize reformed courses. While the features articulated below by evaluators are clearly informed by a shared set of underlying goals, diversity was apparent in that each project is seeking to achieve a particular subset of these features. The discussion assumed that reformed courses that exhibited these characteristics would help students learn science and mathematics concepts.

<u>Philosophy</u>
Teachers and students are seen as joint learners
The link between content and method is made explicit
Collaborative activities foster meta-cognition about teaching and learning
Reform processes promote a positive attitude toward mathematics and/or sciences
Courses are evaluated by multiple methods
Students use data to justify options

<u>Content</u>
Curriculum integrates content across the sciences and mathematics
Content is aligned with standards-based curriculum in science and mathematics
Content represents diversity of opinion
Content is culturally sensitive
Courses are created by multiple input across campuses

<u>Learning</u>
Students experience ambiguity as a result of learning
Reform instills an interest in life long learning in mathematics and science
Course instills confidence
Course fosters active learning
Discovery learning is emphasized as appropriate
Content and process learning are sought
Learning is problem-based
Instruction focuses on an understanding a few central concepts
Students are learning from one another

<u>Instruction</u>
Teachers are engaged with students on individual basis or in small groups
Teachers do not emphasize lecture
Teachers employ appropriate use of technology

Faculty model good instruction
Assessment informs teaching decisions
Problem-solving approach is used
No cookbook labs used
Faculty stress connections with real life experiences

The evaluators shared information and experiences pertaining to the evaluation methods they are using to assess how well the courses being developed and taught by CETP faculties are meeting the goals for their reformed courses. Evaluators noted that it was unrealistic to expect that measurable student outcome changes would occur the first year in newly reformed courses. They needed a "shakedown" period of at least a year. Moreover, one evaluator discussed the difficulty of changing "evaluative" student attitudes in just one course and recommended that student attitude impact be assessed after a series of reform courses.

Course Assessments: Participants mentioned several methods for assessing student learning of course content, including use of student portfolios, final projects and exams. These methods help identify whether, and to what degree, curricular reforms have been implemented. The content of major course exams, for instance, is a good, unobtrusive indicator of the extent of course reform. They can also be used to help identify the specific impacts of the CETP project in settings where multiple reform initiatives are occurring. In order to help establish changes attributable to the program, some of the projects have conducted baseline assessments of mathematics or science content knowledge for students entering teacher preparation programs.

Commonly used course assessment methods can be flawed in various ways. They may not be effective for measuring students' depth of understanding nor suitable for comparative studies. The evaluators considered at some length the question of whether reformed courses should use traditional classroom assessments that may not be aligned with new course goals and learning strategies. Project evaluators felt that using existing instruments may be easier and more reliable, but existing instruments generally do not provide good measures of, or adequately represent the impact of, the project. At the same time, they acknowledged that faculty may have little choice but to use the traditional assessment instruments because it is difficult for them to find or create new instruments that are appropriate to their reform courses or that other faculty members trust.

Focus Groups: The evaluators who have used focus groups explained that this method is valuable for clarifying key issues (such as ways that reformed courses compare with non-reformed courses), identifying opinions and attitudes, developing consensus and determining the strength of feeling about particular issues. They are particularly useful in that they allow students to learn from and build on what the others are discussing. The CETP evaluators favor relatively small focus groups of 5 to 10, and generally find that groups work better when the participants have fairly similar backgrounds. Multiple groups often are needed to ensure sampling diversity. In addition, the evaluators stressed the importance of designing focus groups in a way that is culturally sensitive to the particular group. It was

noted that focus groups have been used by various CETPs to identify key issues that can, on the one hand, be explored in greater depth in interviews, or on the other hand, be verified through surveys. Surveys designed to determine the degree to which large numbers of respondents share the focus group themes.

Observation Data: Some of the evaluators advocated the use of observation for obtaining high quality data on classroom process and behavior. If implemented well, observation methods can be used to focus on students' demonstrations of understanding. The *Rocky Mountain Teacher Education Collaborative* reported involving faculty in action research, such as having a faculty member teach two sections of the same course, one section reformed and the other traditional, and using graduate students to observe differences in student responses.

The Philadelphia Consortium videotaped CETP-infused students and other students without CETP exposure. They analyzed the differences between the two groups using coded data from a behavioral scale they developed consistent with Third International Mathematics and Science Study (TIMSS).

While many of the project evaluators recognize the value of observation, they use this method sparingly because of its high time and training costs. Observers must be carefully trained to capture the complexity and richness of a situation and also to be able to recognize their own biases. Moreover, observation as an assessment approach in the Collaboratives entails the logistical difficulty of visiting multiple sites, sometimes at great distances from each other.

Student Surveys: The meeting participants reviewed commonly understood points about survey methods, namely that surveys are optimal for gathering consistent data on large numbers of faculty and students at fairly low cost, compared with interview or focus group methods. They can be used to gather both qualitative and quantitative data on both reformed courses and comparison courses. The *San Francisco Bay Collaborative (MASTEP)* reported measuring attitudes about teaching with a pre/post student survey which focused on the impact of reforms, on attitude change towards mathematics and science teaching as a profession, as well as evidence of changes in faculty instructional practices. Several projects including the El Paso Partnership for Excellence in Teacher Education report working collaboratively with other CETP projects on student attitudinal surveys.

Drawbacks to surveys also were discussed. For example, evaluators must take into account that surveys are designed to measure preconceived outcomes, and thus are likely to miss important unanticipated processes and outcomes associated with the CETP activities. A problem encountered by a number of evaluators is that some students are given the same survey numerous times if they participate in more than one reformed course. Another problem noted is that surveys are subject to multiple interpretations, depending on how respondents read the questions.

The *Arizona Collaborative for Excellence in Teacher Preparation* reported creating a somewhat unique student survey device in the form of a listserv for students in reform courses. This was an avenue for "true conversation" among students about the reforms. This listserv was used by evaluators for continuous feedback

about the project and can be used to look at longitudinal change. Evaluators found that conversations initiated by faculty are often about content, while conversations initiated by students are often about pedagogy. Two limitations of listservs are that only some students engage, and that the conversations may not be representative of important issues.

Course Reform Standardization: Beyond using measures to assess student learning, Collaboratives frequently need to have a systematic way of assessing the degree to which courses are reformed. The *Montana STEP Project* reported a useful device for helping to standardize course reform information. They created a catalogue presenting data about all reformed courses (location, teacher, times, focus) using input from faculty. It created a checklist of indicators of reform for faculty about their courses, but also used multiple data sources, such as interviews with faculty and students, classroom observation, and student surveys.

Faculty as Evaluators: The *New York Collaborative for Excellence in Teacher Preparation* developed a synergistic technique for promoting course reform. They asked faculty who teach reformed courses to conduct case studies of courses taught by faculty members in their same field but at a different institutions, thereby engaging faculty in reflection and fostering cross-institutional collaboration.

Resources

Evaluators noted that their evaluations are not able to address all the components within their Collaboratives equally. A tension exists between the resources available and the number of project components to be evaluated. Therefore, CETP projects must set priorities regarding evaluation efforts. Projects have focused a large proportion of their evaluation effort historically on curricular and course reform. Project evaluators' resources are already too small for a comprehensive evaluation of the curricular reform sub-system.

Many project evaluators are trying to stretch their resources in order to provide their PI's with evaluation information on other sub-systems, such as longitudinal data, mentoring programs for first year teachers, under-represented student recruitment, faculty development programs, or school outreach efforts. Almost out of the evaluators' reach is the capacity to provide substantive feedback data on the relationships and alignment among these sub-systems. Even with effective use of resources within institutions in their Collaborative to avoid duplicating efforts, evaluators are hard pressed to evaluate the Collaborative's multiple impacts. Project evaluators would like to see the percentage of funds allocated to evaluation increased, given the quantity and quality of work required.

Some projects reported that they have been very resourceful in terms of using graduate student help and "contributed" faculty time and departmental infrastructure. Others pool funds from other NSF programs or if they are in their fifth year, apply for three more years of funding. Opportunities like this meeting, where efforts of one project evaluator can be shared with others, also help to stretch the resources. There are also opportunities to obtain extra funds for special evaluation efforts through REC.

Future Collaborations

This meeting was the first step on a path of improved collaboration among CETP project evaluators. Activities that this group identified as important next steps principally relate to enhancing the quantity and quality of collaboration among projects and between the projects and the program. In addition to the suggestions for cross-CETP activity presented throughout this report, the participants agreed during the final session to initiate and support the following activities:

1. Each project will develop an overview of their Collaborative that will include a narrative component and matrix and that will describe the project's mission, organizational structure, and evaluation plan.

2. Project evaluators and program administration will create a CETP website that links the home pages of each project. Each project will put its evaluation plan and instruments it uses on its website.

3. Evaluators from San Francisco CETP have created an email listserv for all project evaluators.

4. All evaluators will continue to look for ways to share instruments, resources and relevant data. Many evaluation instruments have already been shared

5. Evaluators will try to continue to meet at the American Education Research Association (AERA) conference and find other forums to share insights. This past year, CETP evaluators presented a panel at the AERA annual meeting.

6. Evaluators will do joint studies and jointly develop instruments when possible. Three CETP Collaboratives conducted a pilot pre-service teacher survey project together.

7. When feasible, evaluators will make site visits together.

Summary of the "Across CETP Project" Corridor

Evaluator's believed that using multiple evaluation strategies helped them understand the various types of project effects. The CETP Evaluators' Meeting provided an opportunity to share information and identify many potential opportunities for collaboration across projects. Besides creating additional lines of communication generally, evaluators identified various potentially valuable joint evaluation initiatives. Even though much diversity occurs across CETP projects, the similarities in purpose and activities do provide real opportunities for future sharing and joint activities.

Terry Woodin
National Science Foundation Lead Program Director,
Collaboratives for Excellence in Teacher Preparation

Welcome. We are delighted to see you here and thank Carl Frantz and Frances Lawrenz for the hard work they put into crafting this workshop. With your help they have designed a How To workshop. What are the problems evaluating NSF Collaborative for Excellence in Teacher Preparation (CETP) projects? How are you solving them? Are there lessons to learn from your colleagues in other systemic projects-other CETP, the chemistry and mathematics initiatives, the Alliance for Minority Preparation (AMP) projects, the State and Urban Systemic Initiatives, the Local Systemic change projects?

This is an exciting time to be working on projects focused on teacher preparation and an exciting time to be an evaluator. Both of these fields are receiving an unprecedented amount of interest from the press, the President and the Congress, and state legislatures. The looming combination of a large potential for teacher retirement and the increasing demand for teachers due to state edicts to reduce class size coupled with the growing population of youngsters entering the schools and the increasing demands placed upon teachers by national and state adopted standards in mathematics and the sciences, translate into a projected need for two million new teachers within the next decade, many of whom are needed in science and mathematics. The results of the Third International Mathematics and Science Study (TIMSS) have also been translated into a need for teachers better prepared in the mathematics and sciences.

A growing call for accountability has increased concern for adequate evaluation of projects that began to meet national needs, not only to answer the demands of the Government Performance and Results Act (GPRA), but also to answer the demands of a number of constituents for objective formative and summative evaluation. Hard working dedicated faculty that are trying to improve their courses need help determining whether they are achieving their goals. Do their students preparing to be teachers understand their discipline better? Are these students better able to understand concepts, deal with developing new information and directions in the field, and design effective mechanisms for helping their own students do likewise? Such information is an asset to the faculty member as a useful tool for recruiting second tier faculty to participate and for helping the institutions and departments involved to determine how and if to institutionalize the changes started. Institutional administrators and state and national policymakers need valid evaluation of outcomes in order to make decisions concerning adoption and adaptation of promising practices and programs. The stakes are high for evaluators of systemic projects. Thus, we welcome the practical suggestions you are developing for proceeding with such evaluations within the various collaboratives.

Currently the CETP projects and the program are evaluated on a number of levels. At the project level, each project has a set of evaluators who report directly to the project. The NSF and the project also receive or collect the following data for evaluation purposes: an Annual Report delineating various aspects of each project (generally including

an evaluation report); a set of demographic data (collected by QRC as a subcontract to WESTAT) indicating what is being done and how many people are being affected; and reports of annual site visits by National Visiting Committees (a set of experts external to the site). As the projects mature the type of evaluation changes from formative to summative, gradually shifting the emphasis to outcomes. In addition to the site-based evaluations detailed above, SRI has been contracted to visit a set of current projects in order to evaluate the overall progress and effects of the CETP program. Such attention is heady, however it also imposes responsibility.

The CETP program is almost five years old. There are now seventeen such sites in the country (including the alumni, those who have completed their five-year programs). The total effect of these projects can be gauged by the following set of statistics for calendar year 1997. The fifteen projects that were operating included 160 institutions of higher education (ranging from two year colleges to research 1 institutions); 541 courses have been affected, 52,400 students have been affected, 676 as NSF Teaching Scholars; and 2,700 faculty and 3,500 teachers participated. While anecdotal evidence can indicate the value of these activities, it remains for effective evaluation to reveal the outcomes of this work. The first set of projects is nearing the end of its fifth year and so should be able to provide information in terms of outcomes. How well are their new teachers doing in their classrooms? How many of the activities are supported under the universities? Have collaborations continued? Have reformed teaching practices continued? What is the net result of the various interventions taken within each project? While questions may be asked in a generic way in order to facilitate a broad outlook across all projects, each project should be able to translate the NSF needs for information into something useful for themselves. The numbers and information collected by NSF should serve the site as well as the NSF.

In addition to supporting design, implementation, dissemination and evaluation of the projects, project based educational research is also encouraged. Research projects should be submitted to the Research in Educational Policy Program (REPP) of the Division of Research, Evaluation, and Communication. As an example, ascribing outcomes to specific programs in sites where a variety of systemic projects are in place is not easy and perhaps should be left to those doing research. However, each project can discover what happens within its area of influence and then cite the many possible interventions that could have contributed to those results.

Along with Carl and Frances, Theresa Smith of Oklahoma and Susan Millar of the National Institute for Science Education are participating in this meeting to help capture your ideas and convert them to a usable document. Theresa is an evaluator for the AMP projects and Susan is an evaluator for the Chemistry Initiative. Representatives of SRI, of the Division of Undergraduate Education and of REC are participating at various times to help facilitate and record discussions. We hope this meeting will help to coalesce your ideas concerning effective mechanisms of evaluation; and, that there will be considerable cross fertilization of ideas during the course of your discussions.

Evaluators:

Arizona Collaborative for Excellence in Preparation of Teachers (ACEPT)
Irene Gomez-Felix
Michael Piburn
Mary Kay Stout
Jeff Turley

Los Angeles Collaborative for Teacher Education (LACTE)
Irene Gomez-Felix
Mary Kay Stout

Louisiana Collaborative for Excellence in Preparation of Teachers (LACEPT)
Carl D. Frantz
Noreen Lackett

Western Massachusetts Teacher Education Collaborative (STEMTEC)
Helen L. Gibson
Eric S. Heller

Montana Teacher Education Preparation (STEP)
Carol Thoresen

Maryland Collaborative in Teacher Preparation (MCTP)
Gilbert Austin
J. Randy McGinnis

New Mexico Collaborative for Excellence in Teacher Preparation (NM-CETP)
Ricardo B. Jacquez
Barbara Kimbell
Keith McNeil

New York Collaborative for Excellence in Teacher Preparation (NY-CETP)
Stephen Pape

Oklahoma Teacher Education Collaborative (OTEC)
Brent Roberts
Robert Sinclair

Oregon Collaborative for Excellence in Preparation of Teachers (OCEPT)
Mary Kinnick
Deborah Morris

Philadelphia Collaborative for Excellence in Teacher Preparation (PHILA-CETP)
Jim Degnan
Tony Lutkus

Rocky Mountain Teacher Education Collaborative (RMTEC)
Gail S. Gliner
Jeffrey A. Gliner
Rose Shaw

San Francisco Bay Math and Science Teacher Education Program (MASTEP)
Gloria Guth
Tania Madfes

University of Texas El Paso-Partnership for Excellence in Teacher Preparation (UTEP-PETE)
Thomas Brady
Connie Kubo Della-Piana

Virginia Urban Center Teacher Preparation Collaborative (VCEPT)
George Bass
Marie Sheckel

Other Participants:

Alliance for Minority Participation (AMP)
Theresa Smith

Learning through Evaluation, Adaptation and Dissemination Center (LEAD)
Jennifer Kushner, Meeting Recorder
Susan Millar, Meeting Recorder

Louisiana Tech University
Debbie Silver, Speaker

National Science Foundation (NSF)
Norman L. Fortenberry, DUE Division Director
Joy Fettheling, DUE
Jack Hehn, Program Consultant, DUE (American Association of Physics Teachers)
Con Katzenmeyer, Senior Program Manager, REC
Terry Woodin, Program Manager, DUE

Stanford Research International (SRI)
Barbara Matson, Presenter
Joan Ruskus, Presenter

University of Minnesota
Frances Lawrenz, Meeting Organizer

WESTAT
Gary Silverstein

Meeting Coordinator
Anne Breaux

<div style="border:1px solid black">

Meeting Goals
1) To increase the effectiveness and consistency of CETP formative and summative evaluations
2) To facilitate collaboration and networking among CETP and related evaluators

</div>

Agenda

1) **Wednesday Evening, February 4**

2) **Welcome (7:00-7:15 p.m.) -Carl Frantz**

    i)   Introduction to meeting, description of overall goals and expected outcomes of the meeting
    ii)   Introduction of participants

    b)   **Dinner (7:15-7:40 p.m.)**

    c)   **Welcoming Remarks -Terry Woodin**

    d)   **Presentation: NSF expectations for CETP summative evaluations (7:45-8:45 p.m.) -Conrad Katzenmeyer**

        i)   A discussion about GPRA, use of QRC collected data, presentation of proposal for revamping the CETP evaluation including expectations for CETP final summative evaluations
        ii)   Questions and answers

        Note: Participants will be asked to sign up for interest areas Wednesday evening and through the day on Thursday for the Thursday night dinner discussions. Each Table will be addressing a topic of importance for evaluators at the Thursday night dinner

3) **Thursday Morning, February 5**

    a)   **Networking breakfast (8:00-9:00 a.m.)**
        Participants will gather to get to know each other better, share information about their own evaluation approaches, and discuss common problems and opportunities

    b)   **NSF expectations for summative evaluation continued (9:00-10:15 a.m.) -Frances Lawrenz**

        i)   Small group discussions of NSF expectations (9:00-9:30 a.m.)
            Groups of 4 or 5 will meet for 30 minutes to discuss Wednesday night presentation and offer feedback and further suggestions regarding feasibility, appropriateness, and implications of NSF expectations
        ii)   Plenary session regarding NSF expectations (9:30-10:15 a.m.)
            Small groups will report concerning NSF report expectations
            Plenary session will discuss findings and arrive at general consensus.
            Topics that need further consideration will be identified

    c)   **Break (10:15-10:30 a.m.)**

    d)   **Assessing Student Learning (10:30 a.m.-12:30 p.m.) -Mary Kinnick**

        i)   **Small group discussions on assessing student learning (10:30-11:15 a.m.)**

            (a) Small groups discuss how their assigned methodology has been or could be used to contribute to the assessment of the CETP project impact on student learning

(b) Each group is assigned to one of the following approaches:
        (1) overall research design
        (2) direct content assessment methods (paper and pencil and authentic assessment)
        (3) focus groups
        (4) surveys of students
        (5) utilizing already existing data
        (6) observations
        (7) follow-up in-service teacher assessment

a) **Plenary session on assessing student learning (11:15 a.m.-12:30 p.m.) -Eric Heller**
All groups report their experience regarding and their assessment of the value of their assigned evaluation approach for the task of assessing CETP impact on student learning

4) **Thursday Afternoon, February 5**

a) **Lunch and luncheon speaker (12:20-2 p.m.) -Debbie Silver**
Inspirational speaker and former state teacher of the year, provides a humorous look at standards-based reform and implications for evaluation

b) **Exemplary Practices (2:00-4:30 p.m.) -Stephen Pape**

i) Small group discussions on exemplary practices (2:00-2:45 p.m.)
Small groups will discuss effective student learning or other evaluation techniques or strategies. Every participant will be asked to provide a one-page summary of one or more such exemplary approaches or strategies to share

iii) Plenary session on exemplary practices (2:50-4:30 p. m.)
Each of nine groups present to the full group and discussion will follow

c) **Networking (4:30-7:00 p.m.)**

5) **Thursday Evening, February 5**

a) **Dinner with previously identified discussion groups (7:00-8:00 p.m.)**

b) **Panel discussion on course reform (8:00-9:00 p.m.) -Eric Heller, Carl Frantz, Susan Millar, Mike Piburn**
How do you know when a course is reformed? Panel discussion relating to defining and assessing course reform

6) **Friday Morning, February 6**

a) **Breakfast (7:45-8:30 a.m.)**
Participants will continue networking discussions, share information about their own evaluation approaches, and discuss common problems and opportunities

b) **Small group discussion on evaluation problems and solutions (8:30-9:115 a.m.) - Carol Thoresen**
"I have a problem…(evaluation in nature, of course)". Small groups discuss problems and potential solution. They identify specific problems to share with the next panel

c) **Panel discussion on evaluation do's and don'ts (9:15-10:15 a.m.) -Carol Thoresen, Jeffrey Gliner, Mary Kay Stout**
Experienced evaluators discuss do's and don'ts of CETP evaluation and respond to specific problem situations selected by the groups

d) **Break (10:15-10:30 a.m.)**

e) **Formative evaluation (10:30-11:45 a.m.) -Connie Kubo Della-Piana**

    i)   Small group discussions (10:30-ll:15 a.m.)
    ii)   Report to the full group (11:15-11:45 a.m.)

f) **Lunch (11:45-12:30 p.m.)**

7) **Friday Afternoon, February 6 ***

a) **Fish bowl discussion of reflections on the meeting and on the emerging themes (12:30-1:45 p.m.) -Theresa Smith, Susan Millar, Jennifer Kushner, Frances Lawrenz, Carl Frantz**

    i)   Whole group observes small group of evaluators reflect on the meeting content in relation to other evaluation efforts and identify emerging themes (12:30-1:00 p.m.)
    ii)   Discussion opened to entire group (1:00-1:45 p.m.)

b) **Evaluation collaboration (1:45-3:00 p.m.) -Frances Lawrenz**

c) **Next steps (3:00-3:30 p.m.) -Carl Frantz**

    i)   Evaluators consider appropriate next steps
    ii)   Concluding remarks

* The Friday agenda was modified by adding Terry Woodin, CETP Program Manager, and Norman Fortenberry, DUE Director. They both spoke to the evaluators about various issues of interest and relevance to them.