# SERVING DIFFERENT MASTERS:
## TITLE VII EVALUATION PRACTICE AND POLICY

VOLUME 1 - FINAL REPORT

1993

Paul Hopstock
Malcolm B. Young
Annette M. Zehler

Prepared for:

Office of Policy and Planning
U.S. Department of Education

By Development Associates, Inc.

The report was prepared pursuant to Contract No. LC 89023001 (David Moguel, Project Officer), U.S. Department of Education, Office of Policy and Planning. The views expressed do not necessarily reflect the position or policy of the Department, and no official endorsement by the Department should be inferred.

# TABLE OF CONTENTS

EXECUTIVE SUMMARY . . . . . iii

I. INTRODUCTION . . . . . 1

A. Background . . . . . 1
B. Primary Sources of Information . . . . . 2
C. Study Reports . . . . . 2

II. CONTEXT: EVALUATION OF EDUCATIONAL PROGRAMS . . . . . 3

A. Purposes and Uses of Evaluation by the U.S. Department of Education . . 3
B. Uses of Evaluation for Local Program Improvement . . . . . 5
C. Federal Versus Local Needs: An Inherent Conflict? . . . . . 7

III. EVALUATION OF THE TITLE VII PROGRAM . . . . . 9

A. Federal Statute and Regulations . . . . . 9
B. The Bilingual Education Evaluation System (BEES) . . . . . 11
C. The Evaluation Assistance Centers (EACs) . . . . . 12
D. Purposes of the Current System . . . . . 12

IV. OPERATION AND OUTCOMES OF THE CURRENT SYSTEM . . . . . 17

A. Purposes and Uses of Evaluation . . . . . 17
B. Evaluation of Implementation Processes and Student Outcomes . . . . . 21
C. Quality and Cost of Title VII Evaluations . . . . . 23
D. Qualifications of Evaluators . . . . . 26
E. Evaluation Assistance . . . . . 27
F. Summary . . . . . 29

V. PROGRAM OPTIONS FOR IMPROVING THE SYSTEM . . . . . 31

A. Monitoring of Project Evaluations . . . . . 31
B. Reconfiguring Evaluation Requirements . . . . . 35
C. Redirecting the Evaluation and Research Agenda . . . . . 37
D. Placing Greater Emphasis on Evaluation . . . . . 40
E. Playing a Role in the Selection of Evaluators . . . . . 41
F. Summary . . . . . 43

REFERENCES . . . . . 44

APPENDICES

1. "A Survey of Project Directors and Evaluators Concerning Local Title VII Evaluation and Improvement Practices"
2. "Summary of the Case Study Findings" (from Case Study Report)

# EXECUTIVE SUMMARY

## Background

In 1989, the U.S. Department of Education contracted with Development Associates, Inc., to conduct a review of local Title VII evaluation and improvement practices. The study focused on local Transitional Bilingual Education and Special Alternative projects, and was designed to describe and assess local evaluation practices and the use of evaluation results. This is the final report on the project. This report summarizes the findings and conclusions of the study, and presents program options for the federal government. There is a separate report which presents the results of 18 case studies of the evaluation systems of local Title VII grantees.

## Methodology

The study consisted of three major activities: (1) a file review of the applications and evaluation reports of a stratified random sample of 200 Title VII projects funded in FY 1989; (2) a mail survey of all project directors and evaluators of all 655 projects funded in that same fiscal year; and (3) case studies of the evaluation systems of 18 local projects. In addition, we conducted interviews with all of the federal Project Officers of local Title VII grants within the Office of Bilingual Education and Minority Languages Affairs (OBEMLA), and with other selected education officials at the state and local levels.

## Findings and Conclusions

### Purposes and Uses of Evaluations

1. The purposes and audiences for Title VII evaluations have not been clearly articulated by the U.S. Department of Education. Indeed, there is disagreement within the Department concerning the purposes local project evaluations are to serve.

2. There has been no clear description of how the evaluation needs of the U.S. Department of Education and of local Title VII projects should be integrated, or of how to prioritize them when they are in conflict.

3. There has been some confusion about whether OBEMLA or the Grants and Contracts Service should monitor receipt of evaluation reports. OBEMLA has recognized this problem and made new efforts to monitor the receipt of reports.

4. The timing of application and report deadlines does not realistically allow for review of evaluation reports as part of the grant renewal application process.

5. Of the possible federal uses of Title VII evaluation reports, few, if any, are being realized under the present system. Reports are not being used systematically by the Department to assess Title VII at either the program or project level. OBEMLA ·ɔject Officers might be expected to do so. However, because of their lack of formal ·g in evaluation and statistics and because of the large number of projects for

which they are responsible, OBEMLA Project Officers are not able to perform overall program analyses or to provide detailed feedback to projects about their evaluations.

6. A majority of project directors believe that evaluations of projects could be more useful to those projects than they presently are.

7. An issue frequently encountered by evaluators is the poor quality of data collection systems used by projects. The lack of available high quality process and outcome data hinders many evaluations. The extent to which the improvements of data collection/testing procedures stimulated by evaluation recommendations helped the projects in general or mostly just the project evaluations could not be determined.

**Evaluation of Implementation Processes and Student Outcomes**

8. Of the five most frequently presented elements in Title VII evaluation reports, three are project descriptive variables (training received by staff, parent involvement activities, and instructional methods used) and two are outcome variables (English reading and language skills and oral English skills). Two of these variables are not among those required by federal regulations.

9. Because of the small amount of funds available for evaluation in most projects, project directors often ask evaluators to focus on outcome evaluation. Thus, there tends to be little involvement by evaluators in the evaluation of project implementation processes.

10. Evaluators are more involved in and interested in achievement test results than are project directors and project staff. Project staff are particularly interested in recommendations for improving the design and operations of the project. If project staff had complete control over evaluation designs, there would be much less time and money spent on standardized testing of students.

11. Project directors and evaluators in most cases appear to believe that outcome data should be presented only for the project as a whole. Most have not given serious thought to disaggregating outcome data by such variables as amount of training received by a student's teacher(s), type of intervention received, or extensiveness of intervention.

**Quality and Cost of Title VII Evaluations**

12. Although the cumulative amount spent on Title VII evaluation is quite large, the amount spent for individual evaluations is small. It is unreasonable to expect a detailed and comprehensive evaluation of a major project for $3,000 to $4,000, and it is also unreasonable to expect a meaningful process evaluation when evaluators only spend 4-5 days on-site per year.

13. The average project reported collecting data on three-quarters of the data elements required by Title VII regulations, and including data on half of them in their 1989-90 evaluation report. Given the amount of funds being spent on the average Title VII evaluation, the requirements for evaluation reports in present Title VII regulations are unrealistic.

14. In general, the quality of Title VII evaluation reports could be described as from "poor" to "adequate." There is considerable room for improvement in the evaluations which are done and the evaluation reports which are submitted.

**Qualifications of Evaluators**

15. In most cases, those performing Title VII evaluations have appropriate qualifications. The average Title VII project was evaluated by an individual with 15 years of experience in evaluating educational programs, and 10 years of experience evaluating Title VII projects.

16. The evaluation reports being prepared for the U.S. Department of Education are almost always organized and written by "third-party evaluators."

17. There is no convincing evidence indicating the superiority of either external evaluators or district evaluation staff. The ability and motivation of the evaluator are more important than that person's organizational affiliation.

18. Evaluators who do multiple Title VII evaluations are more likely to meet the mandates of Title VII evaluation regulations. They are not more likely, however, to meet the needs of local projects.

**Evaluation Assistance**

19. The Evaluation Assistance Centers (EACs) are providing useful services to many Title VII projects, but the breadth of their services could be wider. Their present mandate does not include the review of Title VII evaluation reports. Thus, they are not completely aware whether and how their ideas and recommendations are being used.

20. The Bilingual Education Evaluation System (BEES) has been a useful document for evaluators, but it is too technical for project directors. If it were revised, updated, and distributed to evaluators, it would be a useful tool for improving the quality of Title VII evaluations. In order to educate project directors on evaluation, a shorter and less technical document would need to be developed.

21. The Bilingual Education Evaluation System and the Evaluation Assistance Centers have influenced the research designs of Title VII projects. Although some evaluations still include inadequate research designs, research designs are generally better than was indicated in meta-evaluations and summaries conducted in the early 1980s.

Program Options for the Federal Government

**Monitoring of Project Evaluations**

1. Establish a centralized system for keeping track of which evaluation reports have been received, for sending acknowledgements of receipt, and for sending reminders to projects which are late in sending reports.

2. Withhold funds for the next year until the evaluation report for the previous year is received (e.g., no funds delivered for 1992-93 until the 1990-91 evaluation is received).

3. Assign to the EACs the task of reading all evaluation reports and providing written feedback on those reports. This would entail providing feedback in a consistent format, including both positive and negative comments on the evaluation and the report.

4. Clarify that the role of OBEMLA Project Officers in reading the evaluation reports is for grant monitoring purposes only, and not for technical review.

5. Require that each evaluation report include a cover sheet which summarizes key project characteristics and project outcomes for the year. The contents of the cover sheet would be determined by OBEMLA to meet the Department's most important information needs.

6. Maintain a computer file which includes data from the evaluation cover sheets, which would be summarized and reported on yearly.

**Reconfiguring Evaluation Requirements**

7. Delete a number of required data elements for Title VII evaluations, including time on specific tasks, attendance, retention, dropout, gifted and talented, postsecondary enrollment, former project participants, and pre-referral evaluation for handicapped or gifted and talented students.

8. Make data requirements concerning student backgrounds and teacher characteristics more specific (e.g., for students, number of years of schooling in U.S. and other countries).

9. Add data requirements concerning teacher training activities, parent involvement activities, and capacity building.

10. Clarify the evaluation regulations for first year projects. Even if no across-year comparisons are possible, first year projects would report baseline data on project students.

**Redirecting the Evaluation and Research Agenda**

11. Encourage projects to present multi-year data on students in evaluation reports.

12. Encourage projects to include meaningful process evaluation as part of their evaluation systems. Through materials provided to new grantees, Management Training Institutes, etc., OBEMLA would suggest ways in which process evaluation could be integrated into ongoing project activities, and also how agreements with evaluators could be structured so that process evaluation was seen as an important part of the process.

13. Develop formats and standards (but not requirements) for process evaluation of Title VII projects. Such formats and standards would include protocols for the observation of classroom activities, teacher in-service training, and parent involvement activities, and models for providing feedback to teachers and administrators.

14. Define questions of interest to the field which could be answered through systematic review and analysis of Title VII evaluation reports. OBEMLA would contract for such reviews and publish periodic reports on the results of this research.

15. Define those research questions which cannot be answered through review and analysis of Title VII evaluations, and develop approaches for addressing those questions. In many cases, this would involve special federal studies with focused data collection at a limited number of sites.

16. Commission a review and analysis paper on the use of alternative assessment techniques in programs serving LEP students. The paper would address the issues of reliability and validity of alternative assessment techniques, and how such measures can be compared across programs (e.g., in Title VII projects).

**Placing Greater Emphasis on Evaluation**

17. Change the formula for grant review scoring so that the evaluation plan receives more weight. The evaluation plan would be worth at least 15 percent of the total score.

18. Define standards for how much of the project budget should be spent on evaluation, and require that a separate and comprehensive evaluation budget be included in each proposal. Evaluation would account for 8-10 percent of the project budget, although the budget for outside evaluation in most cases would be less (3-7 percent). The evaluation budget would include time/resources for data collection and data file management by either project staff or others.

19. Develop an updated version of the Bilingual Education Evaluation System (BEES) after publication of the next set of Title VII regulations, with distribution to all grantees. OBEMLA would urge grantees to share the updated version of BEES with their evaluators.

**Playing a Role in the Selection of Evaluators**

20. Develop and publish standards for evaluators of Title VII projects. Projects would be strongly encouraged to select evaluators who meet those standards, and would be required to justify selection of an evaluator who does not meet the standards.

21. Develop a roster of persons who have evaluated Title VII projects. This roster would include ratings and comments about the effectiveness, strengths, and weaknesses of the evaluator. This roster would be made available to prospective, new, and ongoing projects.

22. Define the conditions under which district employees may serve as evaluators for Title VII projects. In no case would a project staff member serve as the primary evaluator of a project. However, other district employees who are qualified and who work in a department which is independent of the project would be allowed to serve as primary evaluators.

# I. INTRODUCTION

## A. Background

Under Title VII of the Elementary and Secondary Education Act (ESEA), local school districts and other agencies are funded to provide special services to students from language minority backgrounds who function with limited English proficiency. Two of the programs under Title VII (known as the Bilingual Education Act) are the Transitional Bilingual Education (TBE) program and the Special Alternative (SA) program.

Local projects funded under the TBE and SA programs (as well as other Title VII, Part A programs) are required under federal legislation and regulation to include evaluations of their activities. The evaluation requirements for such projects are described in Sections 500.50 and 500.51 of the regulations. The details of those regulations are described in Chapter III of this report.

In 1989, the U.S. Department of Education (ED) contracted with Development Associates, Inc., to conduct a study entitled "A Review of Local Title VII Evaluation and Improvement Practices." This study focused on TBE and SA projects, and was defined as having nine primary objectives:

- to provide a detailed description of the current evaluation practices of Title VII projects;
- to examine how evaluation results are used by grantees;
- to determine if evaluations being performed are consistent with federal legislation and regulation;
- to assess the comprehensiveness and quality of evaluation methods and processes being used;
- to examine the qualifications of individuals performing local evaluations, and to determine the relationship between evaluator qualifications and quality and utility of evaluations;
- to examine the use and perceived usefulness of resources provided by ED to improve evaluations (i.e., the Bilingual Education Evaluation System (BEES) and the Evaluation Assistance Centers (EACs));
- to determine if the quality and usefulness of Title VII evaluations have improved in the past ten years;
- to identify major problems in federal policy, local practices, evaluator qualifications, training materials, or technical assistance that limit the quality and utility of evaluations; and

- to make recommendations concerning how ED and others can improve the evaluation process and make better use of evaluation results.

B. Primary Sources of Information

The study consisted of three major data collection activities: (1) a document review of the evaluation plans and evaluation reports of 200 TBE and SA projects funded in Fiscal Year 1989 (FY89); (2) mail surveys of all project directors and project evaluators of projects funded in FY89; and (3) case studies of the evaluations of 18 projects funded in FY89. In addition, we conducted interviews with program officers in the Office of Bilingual Education and Minority Languages Affairs (OBEMLA) and with selected officials in state and local education agencies.

The results of the document review were previously presented in a report entitled "A Review of Local Title VII Project Evaluation Plans and Evaluation Reports" (Hopstock & Young, 1990). The results of the interviews with OBEMLA staff were previously presented in a report entitled "Summary of Interviews With OBEMLA Program Staff Regarding Project Evaluations" (Development Associates, Inc., 1990). This report presents the results of the mail surveys and case studies, summarizes the findings and conclusions of the overall study, and presents options for program improvement.

C. Study Reports

There are two reports on the project. This is the final report. It includes a discussion of the uses of evaluation by the U.S. Department of Education and by local programs, a description of Title VII evaluation requirements, a summary of the overall study findings and conclusions, and a set of program options for improving the Title VII evaluation system. As an appendix, the detailed results of the mail surveys of project directors and project evaluators are presented. The appendix also includes a summary of the 18 case studies that are described in the Case Study Report.

In the Case Study Report, we present the results of the 18 case studies of evaluations of Title VII projects. The Case Study Report includes a description of the methodology of the case studies as well as a chapter summarizing the results from the case studies.

This Final Report summarizes the study findings, and is designed for federal, state, and local policy makers and school officials who are interested in how Title VII evaluations might be made to be more effective and useful. The detailed survey results are included in the Final Report so that readers can gain additional insights from project directors and evaluators if they so wish. The Case Study Report is primarily designed for individuals who wish to know the detailed results of our case studies. We believe that there is a wealth of "fascinating detail" in this report; thus, even those who are not interested in the details of individual projects may wish to glance through the case studies.

# II. CONTEXT: EVALUATION OF EDUCATIONAL PROGRAMS

In this chapter, we provide a background and context for the discussion of Title VII project evaluations to be presented in Chapters III, IV, and V. All evaluations have multiple audiences and multiple purposes, and in this chapter we discuss the perspectives and needs of the two major audiences for Title VII evaluations: the U.S. Department of Education and the local programs within which Title VII projects are implemented. At the end of the chapter, we discuss the relationship between federal and local needs, and describe what may be an inherent tension between the two.

## A. Purposes and Uses of Evaluation by the U.S. Department of Education

The federal use of evaluation data concerning educational programs has expanded considerably over the past 25 years. Title I of the Elementary and Secondary Education Act of 1965 represented the first major piece of social legislation that included a mandate for project reporting (McLaughlin, 1975). The requirement for evaluation was added largely due to the efforts of Senator Robert Kennedy, who insisted that there should be some accountability incorporated within the program. The notion of evaluation was controversial, and the 1965 legislation was passed with the evaluation requirement stated in very general language. Thus, state and local school systems were allowed considerable room for interpretation and discretion.

The evaluation requirement was seen by Senator Kennedy as having two purposes: (1) to ensure that the funds were being used to address the needs of disadvantaged children; and (2) to provide information that would empower parents and communities to push for better education for their children. Others, such as Commissioner of Education Keppel, saw the use of information on programs and their effectiveness as a means of upgrading schools. The Commissioner thought that performance comparisons in evaluations could be used to "needle a lot of schools" into improving their performance. Staff in the Office of Education and Department of Health, Education, and Welfare welcomed the opportunity to have additional information about the programs, the populations served, and the educational strategies used. The new Secretary of HEW, William Gorham, promoted the evaluation requirement as a means of finding out "what works" as a first step to promoting the dissemination of effective practices (McLaughlin, 1975).

Thus there were several different viewpoints regarding the purposes of the evaluation requirement. One underlying similarity of these, however, was the expectation of reform and the view that evaluation was central to the development of change. However, there was also a common assumption that evaluation activities, once required, would be "self-executing," i.e., that the requirement would in itself generate objective, reliable, and useful reports, and that findings would be used as the basis of decision making.

Such a result did not occur, however. There was not widespread support for evaluation at the local level, and there were a number of objections to the evaluation requirements. There was also a concern that federal requirements for reporting of outputs would eventually lead to more federal control over schooling (McLaughlin, 1975). The reports from state education agencies to the federal level were not the critical self-assessments that were envisioned.

It was in this context that the evaluation requirements under the original Bilingual Education Act were written. These requirements were similar in basic content to those of Title I. Although not as expressly stated, most of the assumptions about the use of Title VII evaluation also paralleled those applied to Title I. Title VII differed from Title I, however, in that it was a demonstration and capacity-building program rather than a continuing service program. The evaluation requirements which were applied did not reflect this basic difference in purpose.

In the 1970's it became clear that the evaluation requirements in federal education legislation were not generating their desired results. Thus, the reauthorization of Title I in 1974 strengthened the collection of information and reporting of data from local grantees. It also required the Office of Education to develop evaluation standards and models for state and local agencies, and required the Office to provide agencies with technical assistance so that comparable data would be available nationwide, exemplary programs could be identified, and evaluation results could be disseminated (Wisler & Anderson, 1979, Barnes & Ginsburg, 1979). The Title I Evaluation and Reporting System (TIERS) was part of that development effort.

Change occurred somewhat more slowly in the Title VII program. It was not until 1977 that substantive regulations regarding the evaluation of Title VII programs were published. These regulations required that: (1) assessment be made of progress in achieving project objectives; (2) descriptions be provided of instruments of measurement in evaluating the performance of participants; and (3) pretest and posttest results on reading tests for all participating children be provided.

With the 1977 regulations in place, local education agencies were struggling with the problems of evaluating their Title VII projects without adequate funding and without the expertise necessary to develop and use meaningful models. A number of reviews of Title VII evaluations were conducted in the late 1970s and early 1980s, and almost without exception they found that the quality of the evaluations was so poor that no meaningful conclusions could be drawn about the Title VII program as a whole. (A summary of these reviews is presented in Tallmadge et al., 1987.) During this period, Congress and the Department of Education made a number of efforts to address those problems. The 1978 amendments to ESEA directed the Secretary of Education to develop "models for the evaluation of such programs as to the progress made by participants therein attaining English language skills" as well as "evaluation and data gathering models which take into account linguistic and cultural differences of the child." This resulted in the Department contracting for the development of program evaluation and data gathering models for bilingual education, which were circulated in draft form in 1982 but were never released by the Department in final form. In 1980, the Education Department promulgated general administration regulations known as EDGAR, which established criteria for judging evaluation components of grant applications. The 1984 amendments to ESEA directed the establishment of the two Evaluation Assistance Centers and, in 1985, the Department of Education contracted for the development of a Bilingual Education Evaluation System. The 1986 regulations for Title VII added considerable detail concerning the content and methods to be used in Title VII evaluations, and only small changes in those regulations have been made since then.

These various changes in legislation and regulation reflected a continuing federal interest in evaluation data concerning the activities of Title VII projects, and an expanding interest in valid and reliable outcome data. What was not in place, however, was a system at the federal, state, or local levels for using the evaluation results to effect program or project improvement.

In 1988, amendments to ESEA reauthorized the Chapter 1 (formerly Title I) program, and strengthened the emphasis on evaluation and local program improvement. The legislation requires that state agencies identify programs that do not show aggregate achievement gains or which do not make substantial progress toward the goals set by the local school district. Those programs that are identified as needing improvement are required to write program improvement plans. If, after one year, improvement has not been sufficient, then the state agency is required to work with the local program to develop a program improvement process that will succeed in raising student achievement (Billig, 1990).

Unlike Chapter 1, Title VII provides a limited number of discretionary capacity-building grants to local districts on a competitive basis, and state education agencies are not directly involved in the process. Thus, Title VII cannot apply a program improvement process similar to that used in Chapter 1. However, there is still considerable federal interest in making better use of evaluation data to generate school reform and local program improvement. The federal purposes of evaluation which were defined in the 1960s (assuring that the money is well spent, identifying effective approaches, stimulating school reform and improvement) have not changed. The legislative purpose of the Title VII discretionary grants to local school districts (Title VII, Part A) is "to assist local education agencies and other eligible grantees in the development and support of instructional programs" for LEP students, and evaluation has long been perceived as one form of needed assistance. What has increased in recent years is the realization that defining evaluation requirements is only the first step in assuring that evaluation data are used for program improvement purposes.

B. Uses of Evaluation for Local Program Improvement

In much the same way that ideas concerning federal use of evaluation have evolved, so have ideas concerning local use of evaluation. For example, the common assumption that objective evaluation data would be generated and used for local program improvement has not always been supported by actual evaluation experiences.

One important purpose of any project evaluation is to examine and assess the implementation and effectiveness of specific project activities in order to make adjustments or changes in those activities. This type of project evaluation is often labelled "process evaluation." The focus of process evaluation includes assessment of interpersonal relationships, logistics, staff performance, adequacy of facilities, choice of specific staff and project interventions, and appropriateness of design. The changes made as a result of process evaluation may involve immediate small adjustments (e.g., a change in how one particular curriculum unit is presented), minor changes in design (e.g., a change in how aides are assigned to classrooms), or major design changes (e.g., dropping the use of ability grouping in classrooms). Within the context of a program such as Title VII in which the length of the special project is limited, such adjustments or changes may be made during the life of the project or possibly even after the project has been completed. Thus, a major

design change suggested as part of a process evaluation of a Title VII project may be implemented by the local school system after the project is completed.

In theory, process evaluation occurs on a continuous basis. At an informal level, whenever a project teacher talks to a project staff member, they may be discussing adjustments to the project activities or project design. More formally, process evaluation refers to a set of activities in which one or more project managers and/or evaluators observe project activities and interact with project staff and/or students in order to define and communicate more effective ways of addressing project goals. Process evaluation can be distinguished from outcome evaluation on the basis of the primary evaluation emphasis. Process evaluation is focused on a continuing series of decisions concerning project or program improvements, while outcome (or product) evaluation is focused on the results of a particular set of project activities (goal achievement) in a particular time period.

A considerable body of literature has been developed about how evaluation results can and should be used for project or program improvement (David et al, 1989; Glickman, 1991; Meier, 1991; Miles & Louis, 1990; O'Neil, 1990). Much of this literature has taken a systems approach, in which the authors have examined decision making in school systems, and have recommended approaches for generating school change. This literature has identified four key factors associated with effective school reform (David et al., 1989):

- Curriculum and instruction must be reformed to promote higher-order thinking by all students;
- Authority and decision making should be decentralized in order to allow schools to make the most educationally important decisions;
- New staff roles must be developed so that teachers may work together to plan and develop school reforms; and
- Accountability systems must clearly associate rewards and incentives to student performance at the skills-building level.

If there is an overriding theme of much of this literature, it is that there must be "ownership" of the change process by as many of the relevant parties (district and school administrators, teachers, and students) as possible. Change must be seen as a natural and inherent part of the education process, so that individuals in the system accept and feel comfortable with new ways of performing their functions (Meier, 1991).

In any school system, there is considerable inertia which works against change. Evaluators who ignore the forces opposing change often find that their recommendations are ignored. In order to effect change, evaluators and interested program staff need to develop an action plan with specific activities, responsible individuals, and timetables as well as political support from a variety of sources.

In addition to inertia, there are a number of other factors at the local level which limit use of evaluation results. One is the often conflicting needs of special project staff versus general school and district staff. The needs of project staff are primarily for data concerning the nature and effectiveness of project activities (process evaluation), while the needs of others in the district are primarily for data concerning the results of the project (outcome evaluation). Thus, the needs of school and district administrators more closely resemble the

needs of federal officials than they do the needs of project staff. An evaluation which is designed primarily to meet the needs of one of these two local "audiences" will not be as useful for the other audience.

Special project staff also may have strong proprietary feelings about the specific interventions of the project, and thus may seek to limit or hide any negative findings. General district or school staff, on the other hand, may not support specific project interventions, and thus may be looking for negative findings to support their views. To speak of local use of evaluation results in a generic sense may thus ignore differences in administrative responsibility as well as political and personal differences which are very important.

Another major factor affecting local use of evaluation results is the perceived purpose of the evaluation. Many evaluations of federally-supported projects are viewed as being "for" the sponsor, so they are not taken very seriously by local staff. If local school officials or project staff do not believe that an evaluation is addressing questions or issues which are relevant to their needs, they are unlikely either to devote significant local resources to the evaluation (e.g., the development of a project data base system) or to examine and use results for local program improvement.

C. Federal Versus Local Needs: An Inherent Conflict?

It is seldom acknowledged in the evaluation of educational programs that the desires and needs of various audiences for the evaluation may be in conflict. Evaluators sometimes walk a very narrow line between meeting the needs of local project staff (in order to get future work, etc.) and meeting the needs of various audiences outside of the project.

Everyone involved claims to want "objectivity." Each may differ, however, in what constitutes a fair and objective assessment of a program. Each person reading an evaluation brings a different set of motives and expectations, so any evaluation is bound to be viewed more positively by some audiences than by others.

The conflict between federal needs and local project needs is particularly strong. In Table 1 we summarize some of the areas in which this conflict exists. We recognize that federal and local project needs often overlap, and the table by its nature exaggerates the differences between federal needs and local project needs. However, many of the differences cited in the table are very real to persons involved in evaluations of federally-supported local programs, and we believe that they should not be overlooked.

In the next chapter, we describe the Title VII evaluation system. In the first three sections of that chapter, we describe the major components of the system. In the fourth section, we discuss the multiple purposes of the current system and how those purposes interrelate. In that section, we continue the theme that federal needs and local needs for Title VII evaluation data may not always coincide.

TABLE 1

Areas of Conflict Between Federal and Local Project Needs in Evaluation

| Issue | Federal Need | Local Project Need |
|---|---|---|
| Primary focus of the evaluation | Overall effectiveness and impact of the project (outcome evaluation) | Effectiveness of particular activities, staff, etc. (process evaluation) |
| Relationship of evaluator to project | An impartial third party | A member of the project "team" |
| Nature of reporting | Occasional formal reports | Continuous informal feedback |
| Nature of student testing | Standardized, norm-referenced tests | Curriculum-based, criterion-referenced tests |
| Which students are tested | All students involved in the project | Students for whom tests are appropriate and relevant (e.g., students at very low levels are not tested) |
| Use of comparison groups | No-treatment control group would be ideal, but some comparison group is a more realistic goal | All testing resources should be focused on project students |
| Timing of testing | Infrequent, large-scale testing | Frequent, small-scale testing |
| Role of classroom observation, observation of teacher training, etc. | Confirmation of project implementation | Assessment and feedback on effectiveness of particular activities |
| Use of student and teacher background data | Description of project context to "outsiders," cross-project comparisons | Data used to focus instructional and training activities on persons in need |
| Amount of detail concerning school context, project activities, etc. | Detail needed in order to understand project and make cross-project comparisons | Little detail needed because information is already known |
| Treatment of positive results | Reliability checks needed before results are disseminated | Results rapidly published to build local support |
| Treatment of negative results | Formal reporting to provide objective assessment of program | Informal feedback so changes can be quickly made |

A grantee's evaluation must provide information on the academic achievement of:

(A) Current participants in the project, who are:
  1. Children who are limited English proficient; and
  2. Children whose language is English; and

(B) Children who were formerly served in the project as limited English proficient, have exited from the program, and are now in English language classrooms.

This information must include:

(A) The amount of time (in years or school months, as appropriate) the participants received instructional services in the project and, as appropriate, in another instructional setting;

(B) The participants' progress in achieving English language proficiency and, for programs of developmental bilingual education, progress in another language; and

(C) The former participants' academic progress in English language classrooms.

Section 500.51

In carrying out the annual evaluation under S500.50, a grantee shall collect information on-

(A) The educational background, needs, and competencies of the limited English proficient persons served by the project;

(B) The specific educational activities undertaken pursuant to the project;

(C) The pedagogical materials, methods, and techniques utilized in the program;

(D) With respect to classroom activities, the relative amounts of instructional time spent with students on specific tasks;

(E) The educational and professional qualifications, including language competencies, of the staff responsible for planning and operating the project;

(F) The specific activities undertaken to improve prereferral evaluation procedures and instruction programs for LEP children who may be handicapped or gifted and talented.

Section 500.52

A grantee shall report to the Secretary annually, the information collected in S500.51 and an evaluation of the overall progress of the project including the extent of educational progress achieved through the project measured, as appropriate, by-

(A) Tests of academic achievement in English language arts and, for programs of developmental bilingual education, second language arts;

(B) Tests of academic achievement in subject matter areas; and

(C) Changes in the rate of student-
  (1) grade-retention;
  (2) dropout;
  (3) absenteeism;
  (4) placement in programs for the gifted and talented; and
  (5) enrollment in postsecondary education institutions.

# III. EVALUATION OF THE TITLE VII PROGRAM

Since 1968, there have been requirements for Title VII projects to conduct evaluations of their activities. The structure and content of those evaluations have been guided by a number of requirements and resources defined and offered by the Department of Education. In this chapter, we briefly describe those requirements and resources to provide a context for our study findings.

The requirements and resources defined and offered by the Department are: (1) the federal statute and regulations relating to evaluation; (2) the Bilingual Education Evaluation System which was developed for Title VII grantees; and (3) the suggestions and recommendations offered by the Evaluation Assistance Centers at the OBEMLA Management Institutes and special training events.

These three sources are described in separate sections which follow. We conclude the chapter with a section which discusses the purposes of Title VII evaluations and how they relate to the various audiences for evaluation.

## A. Federal Statute and Regulations

The evaluation requirements for Title VII, Part A projects are defined in Section 7033 of the Bilingual Education Act. Those requirements are expanded into regulations in Sections 500.50, 500.51, and 500.52 of the federal code (34 CFR). They include the following:

### Section 500.50

A grantee's evaluation design must include a measure of the educational progress of project participants when measured against an appropriate nonproject comparison group.

The evaluation results must be computed so that the conclusions apply to the persons, schools, or agencies served by the projects.

The evaluation instruments used must consistently and accurately measure progress toward accomplishing the objectives of the project, and must be appropriate considering factors such as the age, grade, language, degree of language fluency and background of the persons served by the project.

The evaluation procedures must minimize error by providing for proper administration of the evaluation instruments, at twelve-month testing intervals, accurate scoring and transcription of results, and the use of analysis and reporting procedures that are appropriate for the data obtained from the evaluation.

The evaluation procedures must provide objective measures of the academic achievement of participants related to English language proficiency, native or second language proficiency (for programs of developmental bilingual education), and other subject matter areas.

Section 501.30 of the regulations describes the basis on which applications for Title VII grants are evaluated. The evaluation plan is worth 8 out of 100 points. According to the regulations, "The Secretary reviews the strength of the evaluation plan and its relationship to the educational goals of the project and the activities conducted to attain those goals."

B. The Bilingual Education Evaluation System (BEES)

In 1985, the Department of Education contracted with RMC Research Corporation to develop an evaluation system for Title VII projects. The Department sponsored this effort because there was widespread concern about the quality of the evaluations which were being performed under Title VII, and about the ability of federal reviewers to draw any conclusions across projects. The Department had earlier sponsored a contract to create Title VII evaluation "models," but these models were never published in final form. The design of the RMC evaluation system was guided by three primary objectives:

- The system was to reflect the sum total of knowledge gained from previous work in bilingual education evaluation.
- The system was to be useful at the local level for purposes of project improvement.
- The system was to be totally responsive to the then-current federal legislation and regulations governing the evaluation of Title VII projects.

The Bilingual Education Evaluation System (BEES) was published in November of 1987. It consisted of a User's Guide in two volumes (Volume I, Recommended Procedures; Volume II, Technical Appendices), plus a brief document entitled "Abbreviated Recommendations for Meeting Title VII Evaluation Requirements."

The BEES is a total evaluation system that involves a process evaluation component, an outcome evaluation component, and procedures for integrating the two. The most innovative element of the system, however, was the gap-reduction design that was recommended for assessing student outcomes.

Volume I of the User's Guide contains ten chapters:

I. Introduction
II. Assuring the Project's Evaluability
III. Planning the Evaluation
IV. Collecting Process Data
V. Selecting/Adapting/Developing Instruments for Assessing Outcome Objectives
VI. Collecting Outcome Data
VII. Implementing the Gap-Reduction Design
VIII. Processing and Analyzing Data
IX. Integrating and Interpreting Results
X. Reporting

Volume II has ten appendices covering such technical topics as classroom observation, functional level testing, test reliability, gap reduction calculations, and correcting for regression.

Since its publication, BEES has been distributed at various times to new project directors and through the Evaluation Assistance Center network. As our survey results show (see Appendix 1, Chapter VI), 38 percent of project directors and 76 percent of project evaluators are aware of BEES, and 58 percent of evaluators report at least some use of the system.

C. The Evaluation Assistance Centers (EACs)

At approximately the same time that the Department contracted for the development of BEES, it also funded two regional Evaluation Assistance Centers (EACs). These centers were designed to assist Title VII grantees and other providers of services to LEP students with materials, training, and technical assistance related to evaluation. Such assistance was previously provided through Evaluation, Dissemination, and Assessment Centers (EDACs) and through Bilingual Education Multifunctional Support Centers (BEMSCs). These other Centers had multiple functions, however, so the EACs were developed to focus specifically on evaluation.

The EACs have assisted OBEMLA by providing presentations at Management Institutes for new projects, they have prepared various materials for projects, and they have provided site-specific assistance through mail and telephone consultation, regional training events, and on-site technical assistance.

The specific content of the assistance which they have provided has, of course, depended on the nature of the request, the person(s) offering assistance, and the mode of assistance. The content of the assistance to Title VII projects has often been guided by the evaluation requirements in the federal regulations and the material in BEES. However, EACs have also focused their training and materials development on emerging issues such as portfolio assessment and other alternative assessment techniques. Our survey findings concerning the operations and effectiveness of the EACs are presented in Appendix 1, Chapter VI.

D. Purposes of the Current System

Evaluations of Title VII projects can serve a variety of purposes at the federal, state, and local levels. There is no explicit statement of purposes in Title VII legislation or regulation, and in discussions with OBEMLA and other Department staff and officials, there was no consensus concerning evaluation purposes. In thinking about Title VII evaluations, we have identified nine major potential purposes which such evaluations could serve. We divide those purposes into three major categories: (a) those that primarily address federal needs; (2) those that primarily address local needs; and (3) those that address both federal and local needs. In describing these purposes, we make a distinction between the national Title VII "program" and the specific "projects" which are implemented by local school districts. These nine purposes are:

**Federal Needs**

(1) To monitor projects to assure that they are being implemented as proposed, and that implementation is within legal guidelines (compliance monitoring);

(2) To examine the project implementation to determine if projects merit further funding (grant funding evaluation);

(3) To judge the overall effectiveness and impact of the program across all projects (program effectiveness assessment);

(4) To answer specific questions about the effectiveness or impact of specific program interventions across projects (meta-analytic research);

(5) To identify and describe exemplary projects and exemplary practices (exemplary project identification);

**Local Needs**

(6) To judge a project's overall effectiveness and impact and to identify and describe the project's more effective and less effective components (project effectiveness assessment);

(7) To provide individual projects with suggestions/recommendations to help them meet more effectively their project goals (project improvement); and

**Federal and Local Needs**

(8) To monitor projects to assure that the activities meet technical standards in terms of effective practice (technical monitoring);

(9) To document project effectiveness and impact in order to develop support for project and program activities in the school system and broader community (project and program promotion).

In Table 2 these nine evaluation purposes are further described in terms of the key evaluation questions related to each. In general, outcome evaluations of projects focus on the sixth and ninth purposes, while process evaluations focus on the seventh and eighth purposes.

TABLE 2

Title VII Evaluation Purposes and Related Evaluation Questions

| Evaluation Purpose | Evaluation Questions |
|---|---|
| **Federal** | |
| 1. Compliance monitoring | 1. Are the intended people being served?<br>2. Are funds being spent appropriately?<br>3. Are the program goals being addressed? |
| 2. Grant funding evaluation | 1. Is the project being implemented as proposed?<br>2. Does the quality of project implementation justify refunding? |
| 3. Program effectiveness assessment | 1. Is there demonstrable progress toward reaching program goals?<br>2. Is the progress toward program goals worth the costs (including opportunity costs)?<br>3. Are the program goals still relevant to the needs? |
| 4. Meta-analytic research | 1. What approaches work best with specific populations?<br>2. What approaches have been shown not to be effective? |
| 5. Exemplary project identification | 1. Which projects demonstrate exemplary practices?<br>2. Which projects or practices can be replicated elsewhere? |
| **Local** | |
| 6. Project effectiveness assessment | 1. How successful was the project in meeting its objectives?<br>2. Which project components were more or less effective?<br>3. What project components should be continued after federal funding ends? |
| 7. Project improvement | 1. Were specific activities implemented as planned?<br>2. Were intended short-term results achieved?<br>3. How can project outcomes be improved? |
| **Federal and Local** | |
| 8. Technical monitoring | 1. Do project activities involve accepted educational practices?<br>2. Are the practices being used appropriate for the populations being served? |
| 9. Project and program promotion | 1. Which project outcomes are most salient and important?<br>2. To whom are those outcomes most important? |

The Title VII evaluation regulations as they are presently written suggest a focus on the first four plus the eighth and ninth of these purposes. The regulations explicitly refer to information which is to be provided to the Secretary of Education, and the contents of the required reports suggest the types of uses which could be made. In communications with projects (through Management Training Institute sessions on evaluation, distribution of BEES, and EAC assistance), it also appears that OBEMLA promotes the sixth and seventh purposes on the list (project effectiveness assessment and project and program improvement). However, there is no clear statement of purpose for Title VII evaluations in the legislation or regulations, and OBEMLA does not clearly indicate how evaluation data are used by federal audiences. Thus, local school officials and local project staff are unclear about why and how evaluation data should be collected and used.

As discussed in Chapter II, there is an inherent tension between evaluation which is primarily conducted for a federal sponsor and evaluation which is conducted for a local project. For the federal sponsor, monitoring and compliance review are important elements of the evaluation process, while for the local project, monitoring and compliance review may actually interfere with the collegial relationships needed to carry out successfully process evaluation. Similarly, the reporting approaches which are most useful to local projects (quick, informal, and focused on specific individuals) have little utility for federal audiences. In Table 3 we rate the importance of the nine purposes of Title VII evaluations for three audiences: federal officials, local school officials, and local project staff. The table demonstrates the divergence of needs of the three groups.

It is important that those designing evaluation systems for federally-supported local projects recognize these conflicts, and to the extent possible, design systems to meet the needs of the various groups. Given the limited resources that have been available for evaluation (the median Title VII project evaluation budget in 1989-90 was $3,500), there has been a natural tendency on the part of those performing evaluations to focus on the legally-mandated needs of the federal sponsor. However, if the needs of local audiences for the evaluation are ignored, the data which are collected may not validly reflect actual project activities and outcomes (because of low motivation for data collection), and important opportunities for project improvement through effective process evaluation may be lost.

In Chapter IV which follows, we present a summary of how Title VII evaluations are conducted, and how they are used by federal and local audiences. In Chapter V, we present program options for increasing the usefulness of Title VII evaluation.

TABLE 3

Importance Of Various Evaluation Purposes For Different Audiences

| Evaluation Purpose | Federal Officials | Local School Officials | Local Project Staff |
|---|---|---|---|
| **Federal** | | | |
| 1. Compliance monitoring | Very important | Moderately important | Not very important |
| 2. Grant funding evaluation | Very important | Not at all important | Not at all important |
| 3. Program effectiveness assessment | Very important | Not at all important | Not at all important |
| 4. Meta-analytic research | Very important | Not at all important | Not very important |
| 5. Exemplary project identification | Moderately important | Not at all important | Not at all important |
| **Local** | | | |
| 6. Project effectiveness assessment | Not very important | Very important | Moderately important |
| 7. Project improvement | Not very important | Moderately important | Very important |
| **Federal and Local** | | | |
| 8. Technical monitoring | Very important | Not very important | Moderately important |
| 9. Project and program promotion | Moderately important | Moderately important | Very important |

# IV. OPERATION AND OUTCOMES OF THE CURRENT SYSTEM

In this chapter, we present our findings and conclusions concerning the Title VII evaluation system for transitional and special alternative projects. Our findings are summarized under five headings:

A. Purposes and Uses of Evaluation
B. Evaluation of Implementation Processes and Student Outcomes
C. Quality and Costs of Title VII Evaluations
D. Qualifications of Evaluators
E. Evaluation Assistance

Under each heading we present selected findings from our surveys, site visits, interviews, and document reviews. The format of presentation is to present the finding and then to follow immediately the finding with our conclusions. More detailed findings are presented in the report on the mail survey included in the Appendix, in Volume 2 of this report, and in the evaluation's interim reports.[1]

## A. Purposes and Uses of Evaluation

In order for an evaluation to be most effective, the participants in and audiences for an evaluation must understand its purposes. Such understanding can guide the design of the evaluation and can help to focus the presentation of results.

---

1. Finding: In interviews, OBEMLA and other Education Department staff, local project directors, and evaluators provided a very wide range of answers when asked about the purposes and audiences of Title VII evaluations. For example, 44 percent of the OBEMLA project officers responsible for local projects thought that evaluation reports were primarily for OBEMLA, 39 percent thought they were primarily for local projects, and 17 percent thought they were intended for both. Interviews with local project directors and evaluators indicated a similar lack of consensus.

   Conclusions: The purposes and audiences for Title VII evaluations have not been clearly articulated by the U.S. Department of Education. Indeed, there is disagreement within the Department concerning the purposes local project evaluations are to serve.

---

1 (a) Hopstock, P. & Young, M. 1990. "A Review of Local Title VII Project Evaluation Plans and Evaluation Reports." Arlington, Va.: Development Associates, Inc. (b) Development Associates, Inc. 1990. "Summary of interviews With OBEMLA Program Staff Regarding Project Evaluations." Arlington, Va.: Development Associates, Inc.

Participants in an evaluation do not need to agree on its purposes, but it is extremely useful if the points of view of the major participants are taken into account. In this way, the participants will be more likely to take the evaluation seriously and to devote energy and resources to its implementation.

---

2. Finding: The present evaluation requirements as defined in the Title VII regulations do not explicitly recognize local needs for evaluation information, or define how such local needs should be met. According to interviews with project directors and evaluators, the Bilingual Education Evaluation System (BEES) and the Evaluation Assistance Centers have helped local projects to define how those needs might be addressed.

   Conclusions: There has been no clear description of how the evaluation needs of the U.S. Department of Education and local Title VII projects should be integrated, or of how to prioritize them when they are in conflict.

---

The legislation and regulations for Title VII prescribe a very comprehensive evaluation for all Title VII projects in every project year. The implication of the legislation and regulations (though it is not clearly stated) is that there will be a very extensive federal review of evaluation findings in order to draw policy-relevant conclusions at both the program and project level.

3. Finding: In our record review, we were able to locate only 54 percent of the evaluation reports which we sought at either OBEMLA or the Grants and Contracts Service.[2] Within OBEMLA, some of the files were in the file room, some were in project officers' offices, and some were split between these locations. We had particular problems locating reports on fourth-year projects (we did not look for fifth-year reports). At the time of the file review, there was no systematic feedback to projects regarding the form or substance of their evaluation reports, nor any indication that reports were missing altogether. Only 58 percent of OBEMLA Project Officers reported having communicated with a local project about a missing report. Since our record review, OBEMLA has instituted new procedures to assure receipt of reports.

Conclusions: There has been some confusion about whether OBEMLA or the Grants and Contracts Service should monitor receipt of evaluation reports. OBEMLA has recognized this problem and made new efforts to monitor the receipt of reports.

---

4. Finding: Project evaluation reports are not due until the December after the project year ends. This deadline is realistic given the need to collect and analyze test score data. Applications for continuation and renewal grants are due in November, and this deadline cannot be made any later without delaying grants. Thus, third-year applications are due before first-year reports, and fourth-year applications are due before second-year reports. This mismatch of application and report deadlines is unfortunate but unavoidable.

Conclusions: The timing of application and report deadlines does not realistically allow for review of evaluation reports as part of the grant renewal application process.

---

[2] Hopstock, P. & Young, M. 1990. "A Review of Local Title VII Evaluation Plans and Evaluation Reports." Arlington, Va.: Development Associates, Inc., p. 5.

---

5. Finding: OBEMLA Project Officers reported spending a limited amount of time reading Title VII evaluations, and also indicated that reports were seldom used in project monitoring (42 percent admitted that they did not use the reports at all). Many also expressed concerns about their ability to provide feedback on the technical aspects of evaluations.

   Conclusions: Of the possible federal uses of Title VII evaluation reports (see Table 3 in Chapter III), few, if any, are being realized under the present system. Reports are not being used systematically by the Department to assess Title VII at either the program or project level. OBEMLA Project Officers might be expected to do so. However, because of their lack of formal training in evaluation and statistics, and because of the large number of projects for which they are responsible, OBEMLA Project Officers are not able to perform overall program analyses or to provide detailed feedback to projects about their evaluations.

---

The usefulness of Title VII evaluations to local projects and schools was another central issue to the study. The study asked a series of questions about the usefulness of various elements of the evaluation process.

---

6. Finding: In the mail survey, 43 percent of project directors rated the evaluation process for their project as "very useful;" 43 percent rated it as "moderately useful;" 11 percent rated it as "of limited use;" and 3 percent rated it as "not at all useful."

   Conclusions: A majority of project directors believe that evaluations of projects could be more useful to those projects than they presently are.

---

---

7. Findings: Two-thirds of the project directors who responded to the survey cited at least one concrete action which had been taken in response to evaluation findings. The most frequent actions taken were to improve data collection/testing procedures (24 percent of projects), to improve staff training (20 percent), and to revise, add, or reorganize instructional activities (18 percent).

   Conclusions: An issue frequently encountered by evaluators is the poor quality of data collection systems used by projects. The lack of available high-quality process and outcome data hinders many evaluations. The extent to which the improvements of data collection/testing procedures cited by project directors helped the projects in general or mostly just the project evaluations could not be determined.

---

B. Evaluation of Implementation Processes and Student Outcomes

The evaluation regulations for Title VII prescribe that a wide range of outcome measures be included in the evaluation reports (e.g., tests of academic achievement in English language arts, changes in the rate of student absenteeism), as well as a wide range of project descriptive variables (e.g., the pedagogical materials, methods, and techniques utilized, the educational and professional qualifications of staff). The regulations do not specifically prescribe that an assessment of the effectiveness of process implementation be included, or that capacity-building (a key purpose of Title VII projects) be addressed.

---

8. Finding: According to project directors, the content areas most often included in evaluation reports were the English reading and language skills of LEP students, project-supported training received by staff, activities to increase parent involvement, the oral English skills of LEP students, and instructional methods used (all present in more than two-thirds of reports).

   Conclusions: Of the five most frequently presented elements, three are project descriptive variables and two are outcome variables. Two of these variables are not among those required by federal regulations.

---

The involvement of project directors and project staff can have a significant effect on the evaluation and how its results are used. If project staff are actively involved in the evaluation, they are more likely to accept the evaluation results as valid and to act on evaluation recommendations. In particular, the project director can set a positive tone for an evaluation through active involvement and interest in evaluation findings.

9. Finding: Project directors were most likely to take part in activities relating to process evaluation (e.g, 87 percent were involved in observing and evaluating project-related classroom activities, and 86 percent were involved in observing and evaluating activities for teachers and parents), while external evaluators were most likely to take part in activities related to outcome evaluation (e.g., 82 percent were involved in performing statistical analyses of test data, drafting evaluation reports, and preparing evaluation conclusions and recommendations). The usefulness of evaluations was rated the highest by project directors when no single person had primary responsibility for the evaluation.

Conclusions: Because of the small amount of funds available for evaluation in most projects, project directors often ask evaluators to focus on outcome evaluation. Thus, there tends to be little outside involvement in the evaluation of project implementation processes.

---

10. Finding: The most useful elements of the evaluation process as rated by project directors were the evaluators' recommendations and the perceptions of parents, teachers, staff, administrators, and students concerning the effectiveness and impact of the project. The most useful elements according to evaluators were achievement test results and the perceptions of parents, teachers, staff administrators, and students.

Conclusions: Evaluators are more involved in and interested in achievement test results than are project directors and project staff. Project staff are particularly interested in recommendations for improving the design and operations of the project. If project staff had complete control over evaluation designs, there would be much less time and money spent on standardized testing of students.

---

11. Finding: Our record review indicated that in most cases, the only ways in which outcome data were disaggegated were by school and grade. In interviews, many project directors reported that results presented in such ways were not useful to the project.

Conclusions: Project directors and evaluators in most cases appear to believe that outcome data should be presented only for the project as a whole. Most have not given serious thought to disaggregating outcome data by such variables as amount of training received by a student's teacher(s), type of intervention received, or extensiveness of intervention.

---

C. Quality and Cost of Title VII Evaluations

Although there is general agreement that federally-supported projects should be evaluated, there is frequently disagreement about the amount of resources that should be devoted to evaluation. There is typically a battle between those (frequently project staff) who say that evaluation resources are largely wasted and should be spent on programs, and those (frequently evaluators and policy makers) who say that programs without effective evaluation are often subject to considerable waste.

In the case of a large federal grant program such as Title VII, it is often difficult to determine how much is being spent on evaluation. Projects usually include evaluation budgets in their grant proposals, but they may actually spend more or less than was proposed, and many of the costs of evaluation (e.g., project staff time for student testing) are frequently not included in the evaluation budget. The total costs of evaluation thus may be greater than simple analysis of project budgets may indicate.

12. Finding: Among the project directors who reported an evaluation budget in our mail survey (81 percent of respondents), the average budget for evaluation in 1989-90 was $4,022. This translates to a total of $1,974,802 for those 491 projects. If this average were applied to all 655 transitional and special alternative projects, the total amount spent on evaluation would project to $2,634,410. In addition, ED is funding two Evaluation Assistance Centers (EACs), the total contracts for which were $1,300,000 in Fiscal Year 1991.

The median project evaluation budget (i.e., half spent more, half spent less) was $3,500. Twenty-eight percent of projects spent less than $3,000 on evaluation. Evaluation budgets represented an average of 2.7 percent of the mean total project budget of $148,840. The median number of hours spent on-site by evaluators was 35 hours. Most evaluator time (a median of 60 hours) was spent analyzing data and preparing reports.

Conclusions: Although the cumulative amount spent on Title VII evaluation is quite large, the amount spent for individual evaluations is small. It is unreasonable to expect a detailed and comprehensive evaluation of a major project for $3,000 to $4,000, and it is also unreasonable to expect a meaningful process evaluation when evaluators spend only 4-5 days on-site per year.

---

As previously described in Chapter III, the evaluation requirements for Title VII projects are very extensive. They require data on at least seven types of outcome variables (at least some of which require multiple measures), and data on at least seven types of project descriptive variables (many of which also require multiple measures). In addition, there are other variables (e.g., teacher training and parent involvement activities) which are very frequently included in evaluation reports because they are such integral parts of the projects.

These data are required to be presented in each yearly evaluation report. Assuming that a project receives full renewal funding, similar data thus are required to be presented five times over the course of the project. In addition, outcome data are to be reported separately for three groups: LEP students in the project, native English speaking students in the project, and students who were defined as LEP and were formerly served by the project.

13. Finding: According to a record review and project directors' reports, the average evaluation contained about half of the elements required by federal regulations. Project directors reported that data were collected on three-quarters of the required elements.

Conclusions: Given the amount of funds being spent on the average Title VII evaluation, the requirements for evaluation reports in present Title VII regulations are unrealistic.

---

There was no significant change in the completeness of reports from 1986 to 1990. First year reports on projects were often very weak. Because many first year projects had barely begun student activities, student outcome data were not available. However, many of these reports did not even include baseline data on students.

Defining the technical quality of an evaluation is considerably more complex than defining its completeness. As part of our record review of applications and evaluation reports, we developed two composite measures of the quality of plans and reports. One of the measures included only elements which were cited in the regulations, and the other included an expanded list. The measures included such elements as whether data were collected on all project students, whether twelve-month testing intervals were used, whether test data were reported in sufficient detail for further analyses, and whether there were reliable measures of process variables.[3]

---

14. Finding: According to the record review, the average Title VII evaluation report contained about half of the quality elements which we had defined. Reports were particularly unlikely to contain data on former participants (13 percent), relate process and outcome variables (15 percent), report nontest data in sufficient detail (28 percent), and report test data in sufficient detail (30 percent).

Conclusions: In general, the quality of Title VII evaluation reports could be described as from "poor" to "adequate." There is considerable room for improvement in the evaluations which are done and the evaluation reports which are submitted.

---

[3] A more complete description of these composites is presented in: Hopstock, P. & Young, M. 1990. "A Review of Local Title VII Evaluation Plans and Evaluation Reports." Arlington, Va.: Development Associates, Inc., p. 21.

D. Qualifications of Evaluators

The qualifications of the evaluator is an extremely important issue because it defines the potential of the evaluation. If the evaluator has an appropriate knowledge base, there is greater probability that the design, implementation, and analysis of the evaluation will meet commonly accepted professional standards. With greater experience, there is also a greater probability that the evaluator will be able to bring program-specific knowledge and perspective to the evaluation. Thus, the background of the evaluator is a key factor in any evaluation.

---

15. Finding: Among those responding to our mail survey, 82 percent of projects reported using an external evaluator for at least some evaluation functions. The persons with primary responsibility for Title VII evaluations (usually external evaluators) reported having extensive credentials (65 percent have a doctoral degree, and an additional 33 percent have a master's degree) and experience (an average of 15 years in evaluating educational programs).

Conclusions: In most cases, those performing Title VII evaluations have appropriate qualifications.

---

---

16. Finding: According to evaluators, in 74 percent of projects, evaluators had primary responsibility for determining the content and outline of evaluation reports. In an additional 11 percent of projects, primary responsibility was shared by two or more persons, often including the evaluator.

Conclusions: The evaluation reports being prepared for the U.S. Department of Education are almost always organized and written by "third-party evaluators."

---

A key question in the study was whether the quality of evaluation was related to evaluator characteristics. In order to address this question we looked at both the organizational affiliation of the evaluator and at the evaluator's background and experience.

---

17. Finding: In our record review, the highest rated evaluation reports in terms of completeness and quality were done by district staff members not associated with the project. According to project directors, external evaluators provided more complete reports and useful evaluation processes than did district staff members.

Conclusions: There is no convincing evidence indicating the superiority of either external evaluators or district evaluation staff. The ability and motivation of the evaluator are more important than that person's organizational affiliation.

---

---

18. Finding: In terms of evaluator characteristics, the most important factor in predicting report completeness was whether the person was doing other Title VII evaluations. Evaluators who performed multiple Title VII evaluations submitted more complete evaluation reports than did those who performed one or two such evaluations. Project directors did not indicate any difference in the usefulness of evaluations based on this factor, however.

Conclusions: Evaluators who do multiple Title VII evaluations are more likely to meet the mandates of Title VII evaluation regulations. They are not more likely, however, to meet the needs of local projects.

---

E. Evaluation Assistance

In the early phases of the Title VII program, there was substantial criticism of the quality of the evaluations which were being performed for the projects (see Chapter II). Analysts who attempted to perform meta-analyses of Title VII evaluations reported that most evaluations were of such poor quality that they could not be so analyzed. Among the criticisms were that appropriate comparison group data were not being collected, and that data collection was not consistent and complete. In response to these criticisms, the Department of Education developed two major initiatives: (1) the funding of two Evaluation Assistance Centers to assist local projects in their evaluations; and (2) the sponsorship of a project in which the Bilingual Education Evaluation System (BEES) was developed. One of the objectives of the present study was to assess the usefulness of those resources.

---

19. Finding: Approximately two-thirds of project directors and evaluators reported having received assistance from an Evaluation Assistance Center (EAC). The most frequent types of assistance were materials sent by mail (42 percent of all respondents), training sessions including staff from more than one project (41 percent), and telephone consultation (35 percent). Of those project directors and evaluators receiving assistance, 93 percent rated it as very useful or moderately useful.

Conclusions: The EACs are providing useful services to many Title VII projects, but the breadth of their services could be wider. Their present mandate does not include the review of Title VII evaluation reports. Thus, they are not completely aware whether and how their ideas and recommendations are being used.

---

---

20. Finding: The evaluators of a majority of Title VII projects reported reading the Bilingual Education Evaluation System (BEES) thoroughly, and using concepts from it in their evaluation. Among project directors, only 10 percent had read it thoroughly.

Conclusions: The BEES has been a useful document for evaluators, but it is too technical for project directors. If it were revised, updated, and distributed to evaluators, it would be a useful tool for improving the quality of Title VII evaluations. In order to educate project directors on evaluation, a shorter and less technical document would need to be developed.

---

One of the specific indicators of quality which was used in this study was the use of the gap reduction design in judging the importance of learning gains. We examined the use of gap reduction because it was suggested in the Bilingual Education Evaluation System as a way of standardizing outcome results across projects. The gap reduction approach was specifically developed because of the poor quality of evaluation designs which were encountered in various reviews of Title VII evaluations in the late 1970s and early 1980s.

---

21. Finding: The most frequent topic of the assistance provided by EACs was research design, including comparison groups and gap reduction (52 percent of projects). The concept or method from BEES which was most often rated as "particularly useful" (33 percent of responses) was gap reduction. Gap reduction analysis was reported by evaluators to be used in approximately one-third of evaluations. In our record review, the term "gap reduction" was used in 18 percent of reports.

Conclusions: The Bilingual Education Evaluation System and the Evaluation Assistance Centers have influenced the research designs of Title VII projects. Although some evaluations still include inadequate research designs, research designs are generally better than was indicated in meta-evaluations and summaries conducted in the early 1980s.

---

## F. Summary

As a way of summarizing our findings concerning the use of evaluation results by federal officials and local project staff, in Table 4 we repeat the possible purposes of Title VII evaluations as defined in Chapter III, define the audience (federal officials or local project staff) for which the purposes are most important, and provide a summary rating of the extent to which those purposes are being met.

Table 4 presents a very sobering picture, especially concerning the federal use of local project evaluations. It is very clear to us that the system could be much more useful both for federal and local audiences. In Chapter V, we present program options for increasing the usefulness of the system.

TABLE 4

Extent To Which Various Evaluation Purposes Are Being Met
(Not at all, Minimally, Somewhat, Moderately well, Very well)

| Evaluation Purpose | Extent To Which Purpose Is Being Met |
|---|---|
| **Federal**<br>1. Compliance monitoring | Somewhat |
| 2. Grant funding evaluation | Minimally |
| 3. Program effectiveness assessment | Minimally |
| 4. Meta-analytic research | Not at all |
| 5. Exemplary project identification | Minimally |
| **Local**<br>6. Project effectiveness assessment | Somewhat |
| 7. Project improvement | Somewhat |
| **Federal and Local**<br>8. Technical monitoring | Minimally |
| 9. Project and program promotion | Somewhat |

# V. PROGRAM OPTIONS FOR IMPROVING THE SYSTEM

In this final chapter, Development Associates presents program options which can be implemented by the Department to improve the Title VII evaluation system. We have organized the program options into five major categories:

(1) improving the monitoring of local project evaluations;
(2) reconfiguring the evaluation requirements for local projects;
(3) redirecting the evaluation and research agenda;
(4) placing greater emphasis on evaluation; and
(5) playing a role in the selection of evaluators.

The format which is used is to first present the program option, and then to discuss the advantages and disadvantages of the option.

## A. Monitoring of Project Evaluations

The federal government has a variety of informational needs concerning the Title VII program and projects. However, many of those needs are not being met. In fact, our findings indicated that in many cases the Department was not aware of whether evaluation reports had been received. The Department also was not systematically reviewing the reports or using the results to develop program summaries.

---

Option 1:

Establish a centralized system for keeping track of which evaluation reports have been received, for sending acknowledgements of receipt, and for sending reminders to projects which are late in sending reports.

---

OBEMLA has already designed a system to address this option, but the system is so new that we are unable to assess its effectiveness. We favor a system in which the acknowledgement of receipt and the follow-up of missing reports is a clerical function performed in a very timely manner. Assessments of the evaluation reports could then be completed at a more reasonable pace and could include positive and negative comments as well as suggestions for improving future evaluations (see Options 3 and 4 below). The advantages of this option are that it increases the likelihood of receipt of evaluation reports and communicates to grantees the federal interest in the evaluation reports. The major disadvantage is the costs of developing and implementing a receipt and control system and of the mailings to grantees.

---

Option 2:

> Withhold funds for the next year until the evaluation report for the previous year is received (e.g., no funds delivered for 1992-93 until the 1990-91 evaluation is received).

---

OBEMLA has also instituted a policy relating to this option, though we are not aware of the extent to which that policy has been implemented. This option is important, because grantees need to know that there will be sanctions if appropriate evaluation reports are not received. The advantage of this option is that it strongly emphasizes the importance of the evaluation requirement in the legislation and regulations. The disadvantage is the time and cost involved in keeping track of missing reports, and the increased coordination required between OBEMLA and the Grants and Contracts Service. However, if funding were contingent simply on receipt of a complete evaluation and not on its quality, the cost in time and resources would be minimal.

---

Option 3:

> Assign to the EACs the task of reading all evaluation reports and providing written feedback on those reports. This would entail providing feedback in a consistent format, including both positive and negative comments on the evaluation and the report.

---

Project directors and evaluators expressed considerable frustration because of the belief that evaluation reports were not being seriously reviewed. We can understand their frustration. Although Project Officers may appear to be the ones who should read and comment on evaluation reports, we believe that EAC staff members are more qualified for this task. As would be expected, the EAC staff have superior credentials to comment on evaluation methods and approaches, and we believe that feedback from EACs would be received in a more accepting manner. We recognize that this option adds a major effort to the scope of work of the EACs, and may need to be postponed until the next EAC procurement process. However, we believe that the task fits in well with other EAC activities. The disadvantage of this option is that it will either require additional financial resources for the EACs or that the EACs may need to decrease their other training and technical resource services.

---

Option 4:

Clarify that the role of OBEMLA Project Officers in reading the evaluation reports is for grant monitoring purposes only, and not for technical review.

---

This option is related to Option 3 above. If EACs had responsibility for the technical review of evaluation reports, OBEMLA Project Officers could focus on whether the intended people were being served and whether the project was being implemented as proposed. By separating the compliance monitoring and technical review functions, the technical review process might be less threatening, and projects might be more open to expressing their uncertainties and concerns to the technical reviewers. The disadvantage to this option is that by separating the compliance monitoring and technical review functions, grantees might not take technical suggestions very seriously. We do not see this as a serious disadvantage, however, because under the current system, very few technical suggestions are being made and acted on.

---

Option 5:

Require that each evaluation report include a cover sheet which summarizes key project characteristics and project outcomes for the year. The contents of the cover sheet would be determined by OBEMLA to meet the Department's most important information needs.

---

The evaluation reports which are submitted to OBEMLA come with a wide range of content outlines and presentation styles. Although Title VII legislation and regulations specify the data which evaluations should contain, there are no specific questions which are required to be answered in the reports, nor are there specified reporting formats or operational definitions. Thus, it is virtually impossible for OBEMLA to answer such questions as how many students were actually served in a particular year, or how many teachers received training. By designing a cover sheet for all evaluation reports to collect data on such questions, OBEMLA could dramatically improve its information base about grantees. It is extremely important that this cover sheet be as simple and self-explanatory as possible, so that the information received is complete and reliable. An example of such a cover sheet is shown in Table 5. The disadvantage of such a cover sheet is that it would have to be approved by the Office of Management and Budget and would add to the total paperwork burden of grantees.

Table 5

Sample Cover Sheet For Title VII Evaluation Reports

Grantee: _____

Grant Number: _____

Number of Schools Served By Grant This Year: _____

Number of Students Served By Grant This Year: LEP-_____

English Proficient-_____

Grade Range of Students Served: _____

Major Language Groups: _____

Number of Teachers Receiving Project-Supported In-Service Training This Year: _____

Average Number of Hours Per Teacher of Project-Supported In-Service Training: _____

Number of Aides Receiving Project-Supported In-Service Training This Year: _____

Average Number of Hours Per Aide of Project-Supported In-Service Training: _____

Average Number of Classroom Aide Hours Per Week Supported By Project Funds (e.g., 7 aides @ 20 hours/aide = 140 hours): _____

| LEP Student Test Score Data Included in Report: Pre- and Post-Test Data Are Presented Concerning: | Yes | No |
|---|---|---|
| English Oral Proficiency | ___ | ___ |
| English Reading | ___ | ___ |
| English Language Arts | ___ | ___ |
| Mathematics | ___ | ___ |
| Social Studies | ___ | ___ |
| Science | ___ | ___ |
| Native Language Arts | ___ | ___ |

---

Option 6:

> Maintain a computer file which includes data from the evaluation cover sheets, which would be summarized and reported on yearly.

---

Assuming that an evaluation cover sheet is developed, it would be extremely important that the data from those cover sheets be summarized and disseminated. By assembling and disseminating this information, ED would: (1) have data available for responding to information requests from Congress and other federal agencies; (2) influence grantees to provide cover sheet and other evaluation data in the future; and (3) generally build a climate of support for Title VII programs. The disadvantage of this option is that it would require new data-base systems and involve additional staff or contractor time for its implementation.

B. Reconfiguring Evaluation Requirements

The Title VII evaluation system could also be improved by reshaping the evaluation requirements. Three of these options involve changes in the supporting legislation, which could be made in the next legislative cycle. The fourth option in this group simply involves clarifying the existing regulations.

---

Option 7:

> Delete a number of required elements for Title VII evaluations, including time on specific tasks, attendance, retention, dropout, gifted and talented, postsecondary enrollment, former project participants, and prereferral evaluation for handicapped or gifted and talented students.

---

Our review of Title VII reports and questionnaire responses from project directors indicates that the number of data elements required in Title VII regulations is too large, and that many of these requirements are being ignored by project evaluators. By decreasing the number of required elements in Title VII evaluations, the coverage of the required elements would be more complete across projects and more comprehensive within projects. In this way, the Department would actually be getting more by asking for less. The disadvantage of deleting these requirements is that the Department would be receiving less information on a systematic basis about certain student outcome variables (though that information at present is not being systematically used).

---

Option 8:

Make data requirements concerning student backgrounds and teacher characteristics more specific (e.g., for students, reports should include the home language and number of years of schooling in U.S. and other countries).

---

In our document review of evaluation reports, it was difficult to decide if reports had addressed the requirements concerning student backgrounds and teacher characteristics. Reports often included some, but not very much, detail. Also, the information included was not consistent across reports. If these requirements were more narrowly defined, it would provide guidance to projects concerning what data to collect and report, and it would facilitate review at the federal level. A disadvantage of this option is that some student background or teacher characteristic data which is presently reported would not be included in evaluation reports. However, because this information is not being consistently obtained, it is of little or no real use at the federal level.

---

Option 9:

Add data requirements concerning teacher training activities, parent involvement activities, and capacity building.

---

The present evaluation regulations do not include requirements for reporting data concerning teacher training, parent involvement, or capacity-building, although these are key aspects of most Title VII projects. In particular, the absence of a data requirement concerning capacity-building is noteworthy, because it is a primary purpose of the Title VII program. The major disadvantage of adding these requirements is the added response burden for projects. However, most existing evaluation reports already include data on teacher training and parent involvement, and data on capacity building is of major importance.

---

Option 10:

Clarify the evaluation regulations for first-year projects. Even if no across-year comparisons are possible, first-year projects would report base-line data on project students.

---

A review of project reports indicated that evaluators were unsure of what should be included in first-year reports. In many cases, no student data were presented. First-year reports can provide very useful baseline information on project students, which can be used to describe student need (the data in applications is often very general) and can be abstracted and integrated into subsequent reports. The main disadvantage to this option is that it might add to the time required for preparing and writing first-year reports.

C. Redirecting the Evaluation and Research Agenda

A third area in which changes can be made in order to improve the Title VII evaluation relates to how the evaluations are done and how they are used. We believe that through its position as a funding source and leader in the field, OBEMLA can encourage and promote more effective evaluation practices and can integrate the various evaluation and research needs. It is important, however, that in this role OBEMLA be seen as a facilitator rather than as an dictatorial authority, because other parts of the bilingual education network also have a great deal to offer.

---

Option 11:

Encourage projects to present multi-year data on students in evaluation reports.

---

Reports which show either longitudinal (same students) or cross-sectional (different groups of students) achievement data across more than two years are particularly useful to federal and local readers in assessing project effectiveness. Such analyses could be encouraged at Management Institutes, through the EACs, and/or through a revised version of the BEES (see Option 19). The addition of such data might make reports more difficult to prepare and understand, but with appropriate guidance, these problems could be alleviated.

---

Option 12:

Encourage projects to include meaningful process evaluation as part of their evaluation systems. Through materials provided to new grantees, Management Institutes, etc., OBEMLA would suggest ways in which process evaluation could be integrated into ongoing project activities, and also how agreements with evaluators could be structured so that process evaluation is seen as an important part of the process.

---

There are various models for performing process evaluation (i.e., examining project activities for the purpose of project improvement). Some models involve activities including only project staff, while others include the inputs (observations, conclusions, and

recommendations) of outside evaluators. We believe that the model for process evaluation which is chosen is much less important than is the level of commitment to examine and improve project activities. Thus, while it is desirable to suggest various approaches, it may be dysfunctional to impose specific requirements in this area.

Based on our site visits, it is clear that a significant amount of process evaluation of Title VII projects is already being done. However, many project directors are unsure of whether or not the Department of Education supports such efforts. We believe that process evaluation is essential to effective project implementation, and that if projects were encouraged and supported in performing it, that significant project improvement could be expected. The major disadvantage of this option is that it might move the evaluation emphasis in some projects away from project outcomes, which is the primary focus of the evaluation regulations.

---

Option 13:

> Develop formats and standards (but not requirements) for process evaluation of Title VII projects. Such formats and standards would include protocols for the observation of classroom activities, teacher in-service training, and parent involvement activities, and models for providing feedback to teachers and administrators.

---

OBEMLA can provide considerable support to local projects in performing process evaluation. By providing suggested formats and standards, OBEMLA will communicate the importance of process evaluation, and will provide a useful service to grantees. However, we believe that the grantee should "be in charge of" the process evaluation and, thus, that the Department of Education should not mandate any particular elements. The disadvantage of this option is that the formats and standards developed by the Department may come to be seen by local projects as requirements rather than suggestions. Should this become the case, it is likely to result in unproductive expenditures of project resources.

---

Option 14:

> Define questions of interest to the field which could be answered through systematic review and analysis of Title VII evaluation reports. OBEMLA would contract for such reviews and publish periodic reports on the results of this research.

---

The amount of money being spent on the evaluation of Title VII projects should be providing a broader knowledge base than is presently the case. Even if the questions are relatively simple ones like "What are the most frequent topics being covered in Title VII supported inservice training?" or "What are the topics or languages of materials being

developed under Title VII grants?", developing answers would meaningfully advance the field. Attempts to answer more complex questions such as "What instructional methods are being used most successfully with older, unschooled LEP students?" would be even more important. As part of its research agenda, OBEMLA would define a set of key research questions, decide which of those questions can be answered through analysis of local Title VII evaluations, and then contract for that review and analysis to be performed, perhaps through its Special Issues Analysis Center (SIAC). The obvious disadvantage of this option is the cost of examining reports to determine which questions can be answered through a review, implementing the reviews, and publishing the results.

---

Option 15:

> Define those research questions which cannot be answered through review and analysis of Title VII evaluations, and develop approaches for addressing those questions. In many cases, this would involve special federal studies with focused data collection at a limited number of sites.

---

There are certain types of research questions which cannot be easily addressed by analysis of local Title VII evaluations (e.g., if a particular intervention approach is not already being used in a consistent manner by a significant number of Title VII projects). In such cases, OBEMLA should design and contract for special studies which are focused on the questions. One cost-efficient approach to conducting such studies might be to ask a group of Title VII projects to use a common research design and common data collection instruments in order to make cross-site comparisons, and to have a contractor perform the cross-site analyses. For example, such comparisons might be made to determine if a similar instructional approach produces similar outcomes in different settings. Cross-site comparisons might also be used to compare different variations of an instructional approach in different settings. The disadvantage of this option is the cost of the special studies.

---

Option 16:

> Commission a review and analysis paper on the use of alternative assessment techniques in programs serving LEP students. The paper would address the issues of reliability and validity of alternative assessment techniques, and how such measures can be compared across programs (e.g., in Title VII projects).

---

Many Title VII projects are dissatisfied with standardized achievement tests and would like to apply alternative methods. However, the evaluation regulations, the Bilingual Education Evaluation System, and the Evaluation Assistance Centers all promote standardized methods in order to assure the validity and reliability of measurement and so that results can be

compared across projects. Under this option, OBEMLA would give serious study to the relevance and usefulness of alternative methods to Title VII projects, and clearly state whether and how the results from such methods meet federal needs. This option presents a difficult challenge to OBEMLA, however, because educational researchers have generally not reached consensus about the usefulness of alternative assessment.

D. Placing Greater Emphasis on Evaluation

A fourth area in which changes could be made to improve the Title VII evaluation system involves the amount of emphasis on evaluation. There are three actions which could be taken to increase that emphasis.

---

Option 17:

> Change the formula for grant review scoring so that the evaluation plan receives more weight. The evaluation plan would be worth at least 15 percent of the total score.

---

The present grant review scoring procedures are such that a proposal with a very poor evaluation plan can still receive a relatively high total score. If evaluation is to be perceived as a key component of any Title VII project, the points for the evaluation plan need to be increased and scorers need to be asked to give low scores to applications with poor or inadequately detailed plans. If this option were adopted, the emphasis on the four other scoring categories would need to be decreased only slightly.

---

Option 18:

> Define standards for how much of the project budget should be spent on evaluation, and require that a separate and comprehensive evaluation budget (including both consultant and project staff costs) be included in each proposal. Evaluation would account for 8-10 percent of the project budget, although the budget for outside consultants in most cases would be less (3-7 percent). The evaluation budget would include time/resources for data collection and data file management by either project staff or others.

---

Without a clearly defined and adequate evaluation budget, the local project is unlikely to gain much meaningful information from a Title VII evaluation. Unless otherwise directed, the first emphasis will tend to be on meeting federal needs, so local needs will be less likely

to be met. There is no commonly accepted formula for how much should be spent on evaluation, but based on our experience, the 8-10 percent range for projects with a total budget of more than $100,000 is generally sufficient to meet both process evaluation and outcome evaluation needs. We believe that the budget for evaluation needs to be viewed broadly to include record-keeping systems and staff time for record-keeping and observation in addition to money for external evaluation. In this way, evaluation activities will be seen as integral parts of the project instead of as "extra work."

Defining the evaluation budget in this way is new to the Title VII program, and there are no data to indicate how much is currently devoted to evaluation defined in this way. We suspect, however, that this option would not substantially increase what is currently being spent on evaluation but would mainly redefine and clarify existing expenses.

Each dollar spent on evaluation is obviously not a dollar spent on programming. For a capacity-building program such as Title VII, however, high quality evaluation data may make the difference between the program being continued or not being continued with local funds.

---

Option 19:

> Develop an updated version of the Bilingual Education Evaluation System (BEES) after publication of the next set of Title VII regulations, with distribution to all grantees. OBEMLA would urge grantees to share the updated version of BEES with their evaluators.

---

Our questionnaires and interviews indicated that BEES has been a useful resource to Title VII evaluators. However, it is rapidly becoming dated, so an updated version is needed. If an updated version were produced, the focus of distribution should be Title VII evaluators, who were the primary users of the previous version. OBEMLA might also consider the development of a simpler and shorter version of BEES specifically designed for project directors and key staff members. The cost of such an updated BEES would likely be considerably less than the cost of developing the original version.

E. Playing a Role in the Selection of Evaluators

The final area in which changes could be made to improve the Title VII evaluation system involves an increased federal role in the selection of evaluators. Without actually selecting the evaluators, OBEMLA could still play an important role.

---

Option 20:

Develop and publish standards for evaluators of Title VII projects. Projects would be strongly encouraged to select evaluators who meet those standards, and would be required to justify selection of an evaluator who does not meet the standards.

---

Our study showed that most evaluators of Title VII projects had appropriate credentials. However, a number of project directors and evaluators said that some evaluations were being done by unqualified persons. The standards would include minimum levels for educational attainment, experience with programs for LEP students, and experience in evaluating educational programs. By publishing standards for evaluators, OBEMLA could signal its commitment to quality evaluation of Title VII projects. The major disadvantage associated with such standards is that they might eliminate from consideration some individuals who are very effective but less credentialed.

---

Option 21:

Develop a roster of persons who have evaluated Title VII projects. This roster would include ratings and comments by past clients about the effectiveness, strengths, and weaknesses of the evaluator. This roster would be made available to prospective, new, and ongoing projects.

---

A number of project directors expressed concern about finding qualified evaluators who were interested in performing Title VII evaluations. Although there is an informal network for identifying such persons, projects could be assisted by the formalization of that network. The roster could be a cooperative effort of OBEMLA and the EACs, and thus the additional costs of producing the roster probably could be kept relatively low. Although it is unclear what information could be legally included in such a government-sponsored listing (e.g., project ratings), at minimum, a listing of prior clients (with telephone numbers) would be included.

---

Option 22:

> Define the conditions under which district employees may serve as evaluators for Title VII projects. In no case would a project staff member serve as the primary evaluator of a project. However, other district employees who are qualified and who work in a department which is independent of the project would be allowed to serve as primary evaluators.

---

There are strong arguments for and against having a district employee serve as the primary evaluator. We believe that OBEMLA should be neutral on this question, as long as the evaluator has appropriate qualifications and is not a regular employee of the project. Regular project employees should not serve as primary evaluators because of the inherent conflict of interest involved in evaluating their own activities.

C. Summary

As our extensive list of program options would indicate, we do not believe that the present Title VII evaluation system is effectively meeting either federal or local project needs. The system is generating volumes of evaluation information, but the data are not being used as well as they might be to generate program and project improvements. The problems are fewer with the persons involved in the system than with the system itself. The evaluation regulations require too much information, inadequate resources are devoted to evaluation, and there is insufficient emphasis on generating information that is useful to the local projects. The program options presented suggest some ways of improving the system.

# REFERENCES


Barnes, R.E., & Ginsberg, A.L. (1979). Relevance of the RMC models for Title I policy concerns. Educational Evaluation and Policy Analysis, 1(2), 7-14.

Billig, S.H. (1990). For the children: A participatory Chapter 1 program improvement process: Paper presented at the Annual Meeting of the American Educational Research Association, Boston, MA.

David, J.L., Purkey, S., & White, P. (1989) Restructuring in Progress: Lessons From Pioneering Districts. Washington, DC: National Governors Association.

Development Associates, Inc. (1990). Summary of interviews with OBEMLA program staff regarding project evaluations. Arlington, VA: Development Associates, Inc.

Glickman, C. (1991). Pretending not to know what we know. Educational Leadership, 48(8).

Hopstock, P., & Young, M. (1990). A review of local Title VII project evaluation plans and evaluation reports. Arlington, VA: Development Associates,Inc.

McLaughlin, M.W. (1975). Evaluation and reform: The Elementary and Secondary Education Act of 1965/ Title I. Cambridge, MA: Ballinger.

Meier, D. (1987). Central Park East: An alternative story. Phi Delta Kappan, June, 1987.

Miles, M.B. & Louis, K.S. (1990). Mustering the will and skill for change. Educational Leadership, 47(8).

O'Neil, J. (1990). Piecing together the restructuring puzzle. Educational Leadership, 47(7).

Tallmadge, G.K., Lam, T.C.M., & Gamel, N.N. (1987). The evaluation of bilingual education programs for language-minority, limited-English-proficient students: A status report with recommendations for future development. Mountain View, CA: RMC Research Corporation.

Wisler, C.E., & Anderson, J.K. (1979). Designing a Title I evaluation system to meet legislative requirements. Educational Evaluation and Policy Analysis, 1(2), 47-55.

# APPENDIX 1

## A Survey of Project Directors and Evaluators Concerning Local Title VII Evaluation and Improvement Practices

DRAFT

# "A SURVEY OF PROJECT DIRECTORS AND EVALUATORS CONCERNING LOCAL TITLE VII EVALUATION AND IMPROVEMENT PRACTICES"

Paul J. Hopstock, Ph.D.

Prepared for:

The U.S. Department of Education
Planning and Evaluation Service

Under Contract No. LC 89023001

Development Associates, Inc.
1730 North Lynn Street
Arlington, VA 22209-2009

1992

# EXECUTIVE SUMMARY

## Background

This report presents the results of a mail survey of project directors and evaluators concerning evaluation practices and use of evaluation by Title VII transitional and special alternative projects. All Title VII projects are required to perform yearly evaluations of their project activities and to report data on a range of outcome variables to the U.S. Department of Education. The study also included a document review and 18 case studies the results of which are presented in separate documents.

## Methodology

Questionnaires were sent to all project directors and project evaluators of transitional and special alternative Title VII projects (N=665). After extensive follow-up, responses were received from 608 project directors and 560 evaluators (93 percent and 85 percent of all potential respondents).

## Summary of Findings

- The three most frequently reported objectives of Title VII projects were improving the oral English skills of LEP students, improving the English reading skills of LEP students, and providing training opportunities for school staff serving LEP students.

- The average Title VII project (as defined by the median) served 245 students in three schools.

- External evaluators (i.e., those not regularly employed by the school district) had primary responsibility for 67 percent of evaluations and had some involvement in at least 82 percent.

- Evaluators for 65 percent of projects had either a Ph.D. or Ed.D.

- The average Title VII project was evaluated by an individual with 15.3 years of experience in evaluating educational programs, and 10.0 years of experience evaluating Title VII projects.

- Where project funds were used for evaluation, the median amount spent was $3,500. Larger projects tended to spend for evaluation.

- Of all Title VII evaluations, 93 percent used norm referenced standardized achievement tests, 72 percent used English oral proficiency tests, 39 percent used criterion referenced achievement tests, and 32 percent used classroom grades as student outcome measures.

- Project directors were most involved in process evaluation activities (observing and evaluating classroom, teacher, and parent activities), but outside evaluators were

most involved in outcome evaluation (summarizing evaluation data, performing statistical analyses, and drafting reports). Project directors and outside evaluators shared the responsibility for presenting evaluation results to project and school staff and school administrators.

- Project staff devoted an average of approximately 80 staff days to evaluation activities. The largest amounts of time were devoted to student testing and collection of existing student records.

- The evaluator at a median project spent 35 hours on site in performing the evaluation and 60 hours analyzing data and preparing reports.

- Although the evaluators of 55 percent of the projects had read the Bilingual Education Evaluation System thoroughly, the directors of only 10 percent of the projects had done so.

- The project directors and evaluators of 65 percent of projects reported receiving some form of assistance from the evaluation assistance centers. The types of assistance which were most frequently received were materials sent by mail, multi-project group training sessions, and telephone consultation.

- The most frequently reported elements in Title VII evaluation reports were English reading and language skills of LEP students at the end of the project year (87 percent of projects), project-supported college or university training received by project staff (87 percent), activities to increase parent involvement (83 percent), English reading and language skills of LEP students at the beginning of the project (80 percent), oral English skills of LEP students at the end of the project year (75 percent), and oral English skills of LEP students at the beginning of the project (75 percent).

- The average project reported collecting data on three-quarters of the data elements required by Title VII regulations and reported including half of the elements in their 1989-90 evaluation report.

- Project directors of 79 percent of projects reported holding meetings or serious discussions with project staff about evaluation results, and project directors of 61 percent reported holding meetings or serious discussions with school administrators.

- Project directors rated the evaluation process as very useful in 43 percent of the cases, moderately useful in 43 percent, and of limited use or not at all useful in 14 percent.

- The three most frequent actions taken in response to evaluation results were to improve data collection or testing procedures (24 percent of all projects), to improve staff training (20 percent), and to add, revise, or reorganize instructional activities (18 percent).

# I. INTRODUCTION

## A. Background

Under Title VII of the Elementary and Secondary Education Act (ESEA), local school districts and other agencies are funded to provide special services to students from language minority backgrounds who function with limited English proficiency. Two of the programs under Title VII (known as the Bilingual Education Act) are the Transitional Bilingual Education (TBE) program and the Special Alternative (SA) program.

Local projects funded under the TBE and SA programs (as well as other Title VII, Part A programs) are required under federal legislation and regulation to include evaluations of their activities. The evaluation requirements for such projects are described in Sections 500.50 and 500.51 of the regulations. The details of those regulations are described in Volume 3, Chapter 3 of this report.

In 1989, the U.S. Department of Education (ED) contracted with Development Associates, Inc., to conduct a study entitled "A Review of Local Title VII Evaluation and Improvement Practices." This study focused on TBE and SA projects, and was defined as having nine primary objectives:

- to provide a detailed description of the current evaluation practices of Title VII projects;
- to examine how evaluation results are used by grantees;
- to determine if evaluations being performed are consistent with federal legislation and regulation;
- to assess the comprehensiveness and quality of evaluation methods and processes being used;
- to examine the qualifications of individuals performing local evaluations and to determine the relationship between evaluator qualifications and quality and utility of evaluations;
- to examine the use and perceived usefulness of resources provided by ED to improve evaluations (i.e., the Bilingual Education Evaluation System (BEES) and the Evaluation Assistance Centers (EACs));
- to determine if the quality and usefulness of Title VII evaluations have improved in the past ten years;
- to identify major problems in federal policy, local practices, evaluator qualifications, training materials, or technical assistance that limit the quality and utility of evaluations; and

- to make recommendations concerning how ED and others can improve the evaluation process and make better use of evaluation results.

The study consisted of three major data collection activities: (1) a document review of the evaluation plans and evaluation reports of 200 TBE and SA projects funded in Fiscal Year 1989 (FY89); (2) mail surveys of all project directors and project evaluators of projects funded in FY89; and (3) case studies of the evaluations of 18 projects funded in FY89. This report presents the results of the mail surveys. The detailed results of the document review are presented in a separate document. The findings and conclusions of the overall study as well as policy options are presented in the study's final report.

B. Contents of This Report

This report presents the results of the mail surveys of project directors and evaluators of Title VII projects. In the next chapter (Chapter II) we describe the methodology of the survey, the nature of the projects included in the survey, the major analytic variables, and some survey definitions. In Chapters III through VIII we describe the results from the surveys. The sequence of presentation is to answer the questions "who," "what," "how," and "why" in that order. Thus, Chapter III describes evaluation staffing, Chapter IV describes the outcome measures used in evaluation, Chapter V describes how evaluations are actually implemented, and Chapter VI describes the use of outside resources to assist in evaluation. The final two chapters describe the reporting and dissemination of evaluation results (Chapter VII) and how evaluation results are used (Chapter VIII). Chapter VII concludes with a summary of recommendations by project directors and evaluators to improve the quality and usefulness of evaluation.

# II. METHODOLOGY

## A. Survey Methodology and Instruments

There were two groups of survey respondents: (1) the project directors of all transitional (i.e., TBE) and special alternative (SA) projects funded under Title VII in FY89; and (2) the evaluators (i.e., the persons who had primary responsibility for the evaluation) for the 1989-90 school year of those same transitional and special alternative projects. There were a total of 655 projects in the population.

Prior to the survey, district superintendents were informed of the study and asked if the contact person for the project in our files was correct. After making some corrections to our list of contact persons, we mailed both the project director and evaluator questionnaires to the project contact person. We asked the contact person to distribute the questionnaires to the appropriate respondents. Included in the package was a postcard on which the contact persons were to record the evaluator's name, address, and telephone number to facilitate follow-up.

We performed very extensive follow-up of the mail questionnaires. Initially, we called project directors and evaluators to remind them to return the questionnaires. After a number of follow-up calls, we requested help from the Office of Bilingual Education and Minority Languages Affairs (OBEMLA) to encourage projects to respond. In a number of cases, we offered to complete the questionnaire by telephone. By the time of the close of data collection (September 3, 1991), we had received a total of 608 useable project director questionnaires and 560 evaluator questionnaires. These numbers reflect response rates of 93 percent and 85 percent, respectively. We received one project director questionnaire after the close of data collection.

We were very pleased with the response rates given the following facts: (1) the respondent burden for the questionnaires was relatively high (approximately 40 minutes per form); (2) many of the respondents had to complete more than one questionnaire (one evaluator completed 24 questionnaires); (3) the questionnaires in some cases requested information about projects which had ended; and (4) in some cases the project director or evaluator for the relevant year (1989-90) was no longer available to answer questions. For external evaluators in particular, there was often no financial or other reward for taking part in the survey. We are thus very appreciative for the contributions of all of our respondents.

## B. Study Definitions and Presentation Style

The survey instruments asked respondents to describe the evaluation activities and outcomes as they related to a particular Title VII project that was identified on the cover of the questionnaire. Most of the questions in the surveys referred to the 1989-90 school year. We requested information about the 1989-90 school year because, by the time of the surveys, all reports and presentations concerning that school year should have been completed.

The survey data that are presented in the remainder of this report are population results. That is, the results are not subject to sampling error, though they are subject to error based on missing data and inaccurate measurement. Thus, the report does not present data on the statistical significance of results. The report does make reference to significant differences among subgroups. When the term "significantly different" is used, it refers to population differences which we believe are large enough to have significance at the policy or program level.

Data presentations throughout this report employ the following conventions:

- All presentations exclude missing data from percentages and means; the number of cases on which those percentages and means are based are clearly presented.

- Results are presented to the nearest whole percentage point or to one decimal place for most scaled values.

- When percentages in a table add to 100 percent, a total column or row is included; when percentages do not add to 100 percent, no total column or row is presented.

- Percentages in a column or row may not add to 100 percent due to rounding error; nonetheless, the total will be labelled "100%" to indicate that all nonmissing responses are included in the total.

C. Major Analytic Variables

Throughout this volume, comparisons are made among different categories of projects. We have chosen three major variables to use in making those comparisons:

- Year of initial funding, which has five subgroups:
  (1) projects initially funded in 1985;
  (2) projects initially funded in 1986;
  (3) projects initially funded in 1987;
  (4) projects initially funded in 1988;
  (5) projects initially funded in 1989.

- Type of project, which has two subgroups:
  (1) transitional projects;
  (2) special alternative projects.

- Size of project, which has four subgroups:
  (1) less than $100,000;
  (2) $100,000 - $149,999;
  (3) $150,000 - $199,999;
  (4) $200,000 or more.

The number of cases in these subgroups varies based on whether the results are from the project director or evaluator survey. Whenever subgroup results are presented, however, the number of nonmissing cases in each subgroup is defined.

D. Description of Projects in the Survey

At the beginning of the project director questionnaire, we asked a series of questions about the nature of the projects supported by Title VII. These questions were intended to provide a context in which to describe the evaluation of those projects.

The first question we asked concerned the number of schools served by the Title VII project. The results are shown in Table 1. The mean number of schools served was 4.6, and the median number was 3.

TABLE 1

Number of Schools Served By Title VII Projects
(N=584)

| Number of Schools | Percentage of Projects |
|---|---|
| 1 | 25% |
| 2 | 20 |
| 3 | 15 |
| 4 | 10 |
| 5 | 8 |
| 6 | 7 |
| 7-9 | 6 |
| 10-19 | 7 |
| 20 or more | 2 |
| Total | 100% |

We also asked about the number of students served by the Title VII project in 1989-90. The results of this question are presented in Table 2. The mean number of students served was 363, and the median was 245.

TABLE 2

Number of Students Served By Title VII Projects
(N=545)

| Number of Students | Percentage of Projects |
|---|---|
| 100 or less | 19% |
| 101-200 | 25 |
| 201-300 | 19 |
| 301-500 | 20 |
| 501-1000 | 12 |
| 1001 or more | 6 |
| Total | 100% |

We also asked if there were any other Title VII projects which were serving (or had served in the previous five years) the same schools served by the project of interest. Of those responding (N=596), 29 percent responded "yes" and 71 percent responded "no."

Table 3 shows the grade levels served by Title VII projects. Projects tend to be focused at the elementary grades, but at least 30 percent of projects are serving students at each of the grade levels from kindergarten to grade 12.

TABLE 3

Grade Levels Served By Title VII Projects
(N=582)

| Grade Level | Percentage of Projects |
|---|---|
| PK | 7% |
| K | 53 |
| 1 | 57 |
| 2 | 57 |
| 3 | 56 |
| 4 | 50 |
| 5 | 50 |
| 6 | 48 |
| 7 | 42 |
| 8 | 42 |
| 9 | 37 |
| 10 | 32 |
| 11 | 31 |
| 12 | 30 |

Finally, we asked project directors to indicate the objectives which were addressed in a significant way by their project. Project directors were told to check only those objectives for which the project should be held accountable. The responses are shown in Table 4. Project directors checked an average of 10.3 of the objectives listed in the table.

TABLE 4

Objectives of Title VII Projects
(N=608)

| Objective | Percentage of Projects |
|---|---|
| Improving the oral English skills of LEP students | 93% |
| Improving the English reading skills of LEP students | 93 |
| Providing training opportunities for school staff serving LEP students | 90 |
| Improving the English writing skills of LEP students | 86 |
| Increasing parent involvement | 83 |
| Improving the cultural awareness of LEP students | 73 |
| Improving the self-concepts of LEP students | 72 |
| Introducing new instructional methods for use with LEP students | 67 |
| Decreasing grade retention, dropout, or absenteeism among LEP students | 65 |
| Introducing or developing new curricula or materials for use with LEP students | 64 |
| Improving math achievement levels of LEP students | 62 |
| Improving science or social studies achievement levels of LEP students | 54 |
| Improving the oral native language skills of LEP students | 46 |
| Improving the native language reading and/or writing skills of LEP students | 40 |
| Increasing high school graduation rates or enrollment in post-secondary education institutions | 29 |
| Other | 14 |

# III. EVALUATION STAFFING

## A. Types of Staff Used

Perhaps the first question to be asked in describing any evaluation is who performed the evaluation. We asked project directors to indicate the person who had the primary responsibility for carrying out the evaluation. The results are presented in Table 5. As the table shows, most evaluations are conducted by external evaluators (i.e., those not regularly employed by the school district). In a significant number of cases, however, the project directors failed to indicate a single person who had primary responsibility but, instead, indicated two or more persons with responsibility.

TABLE 5

Person With Primary Responsibility for Evaluation
(N=604)

| Type of Person | Percentage |
|---|---|
| External evaluator | 67% |
| A district testing/evaluation staff member | 10 |
| Project director | 8 |
| A project staff member other than the project director | 1 |
| Another district employee | 0 |
| More than one of the above | 14 |
| Total | 100% |

We looked to see if the type of evaluator used varied for different types of projects. We found no major differences, however, based on project type (transitional vs. special alternative), year of initial funding, or size of grant.

We also asked evaluators if they had any staff helping them with the evaluation. Of those responding (N=555), 76 percent said that they had at least one staff member helping them with the evaluation. Among those reporting staff assistance, the mean number of assistants reported was 2.9 (1.6 professionals and 1.3 clerical/support staff members).

## B. Selection of the Evaluator

We asked project directors to describe the manner in which evaluators were selected for the project. The results of this question are presented in Table 6. The selection method most often employed was an informal competitive process although a variety of other methods were also frequently used.

TABLE 6

Method of Selection for Evaluators
(N=586)

| Method of Selection | Percentage |
|---|---|
| The evaluator was hired through an informal competitive process (i.e., using the informed judgment of the project director or other district staff member). | 32% |
| The evaluator was hired by the project or district with no competition (e.g., based on personal knowledge). | 23 |
| The evaluator was hired through a formal competitive process (i.e., selection criteria, judging of candidates, etc.). | 21 |
| There was no selection process; the primary evaluator is an assigned district employee. | 16 |
| The primary evaluator was selected by the project from a list of district employees. | 1 |
| Other method or combination of the above. | 6 |
| Total | 100% |

Most evaluation handbooks and manuals suggest that it is important that evaluators be actively involved early in the project. We, therefore, asked evaluators when they first became involved in evaluating the project. The results are presented in Table 7. Slightly less than half of the evaluators became involved before the project actually began operations. However, 19 percent had begun their involvement after the first year of the project.

TABLE 7

When Evaluators Became Involved in the Project
(N=556)

| Time of Initial Involvement | Percentage |
|---|---|
| When the initial application was being developed | 38% |
| After the initial application, but before the project began | 10 |
| Within the first three months of the project | 20 |
| After the first three months, but within the first year | 13 |
| After the first year of the project | 19 |
| Total | 100% |

C. Evaluator Background, Training, and Experience

We asked persons who evaluated Title VII projects a series of questions about their background and experience. One of the issues in describing the results is that a number of evaluators were responsible for more than one project. Thus, our choice was whether to use the project or the individual evaluator as the unit of analysis. Because the focus of the study was on projects rather than evaluators, we decided to have the evaluator complete one questionnaire per project. Thus the results in most of this section use the project as the unit of analysis. At the end of the section, however, we summarize the results using the evaluator as the unit of analysis.

First, we asked evaluators to indicate their highest academic degree. Of the 553 projects, 46 percent had an evaluator with a Ph.D., 19 percent had an evaluator with a Ed.D., 33 percent had an evaluator with a master's degree, 1 percent with a bachelor's degree, and 1 percent with some other type of degree.

We also asked about the major field in which the respondent got the highest degree. The results are presented in Table 8. The most common degree area was educational administration.

We asked evaluators how many college- or university-sponsored courses they had taken which were related to four topic areas: (a) methods for evaluating educational programs; (b) measurement of educational progress; (c) data analysis of educational data; and (d) approaches to educating limited English proficient students. Table 9 summarizes those data.

We also asked evaluators how many hours of other professional training they had received in the same four areas. The results of that question are presented in Table 10.

TABLE 8

Major Field of Study of Evaluators
(N=551)

| Major Field | Percentage |
|---|---|
| Educational administration | 31% |
| Psychology/ educational psychology | 22 |
| Education/ curriculum and instruction | 18 |
| Evaluation and research/ statistics | 9 |
| Linguistics/ language/ communication | 9 |
| Educational measurement | 3 |
| Anthropology/ sociology | 2 |
| Social science | 2 |
| Others (multicultural education, social psychology, etc.) | 4 |
| Total | 100% |

TABLE 9

College and University Coursework Taken By Evaluators
(N=509)

| Topic Area | Number of Courses | | | | | Mean |
|---|---|---|---|---|---|---|
| | 0 | 1-2 | 3-5 | 6+ | Total | Number |
| Methods for evaluating educational programs | 12% | 44 | 26 | 18 | 100% | 3.73 |
| Measurement of educational progress | 7% | 42 | 36 | 15 | 100% | 3.38 |
| Data analysis of educational data | 8% | 30 | 41 | 21 | 100% | 3.94 |
| Approaches to educating limited English proficient students | 30% | 29 | 16 | 26 | 100% | 3.85 |

TABLE 10

Hours of Professional Training Received By Evaluators
(N=453)

| Topic Area | Hours of Training | | | | | Mean |
|---|---|---|---|---|---|---|
| | 0 | 1-20 | 21-50 | 51+ | Total | Number |
| Methods for evaluating educational programs | 12% | 34 | 24 | 30 | 100% | 66.1 |
| Measurement of educational progress | 17% | 38 | 24 | 21 | 100% | 45.9 |
| Data analysis of educational data | 17% | 37 | 22 | 23 | 100% | 47.0 |
| Approaches to educating limited English proficient students | 10% | 26 | 26 | 37 | 100% | 71.8 |

We asked evaluators to describe their experience in evaluating educational programs in general and Title VII projects in particular. Table 11 summarizes the data concerning years of experience. The median number of years evaluating educational programs was 16, and the median number of years evaluating Title VII was 9.

TABLE 11

Years of Evaluation Experience of Evaluators
(N=553)

| Years of Experience | Educational Programs | Title VII Projects |
|---|---|---|
| 0-3 | 8% | 20% |
| 4-9 | 15 | 31 |
| 10-19 | 41 | 35 |
| 20+ | 36 | 15 |
| Total | 100% | 100% |
| Mean Years of Experience | 15.3 | 10.0 |
| Median Years of Experience | 16 | 9 |

Finally, we asked whether the evaluator had been involved in evaluating more than Title VII project. Of the 557 projects, 86 percent were being evaluated by a person who had done or was doing other Title VII evaluations. Respondents reported having evaluated up to 200 Title VII projects in their careers and up to 44 projects at the time of the study.

As an alternative form of analysis, we looked at many of these same variables using the evaluator rather than the project as the unit of analysis. Based on the project directors' responses, we identified 275 separate individuals/organizations who had conducted Title VII evaluations. Table 12 shows the number of evaluations that those evaluators had performed. Although 65 percent of evaluators were reported to be evaluating only one project, 3 percent were reported to be evaluating 10 or more projects, and one was reported to be evaluating 24 projects.

TABLE 12

Number of Evaluations Performed By Individual Evaluators
(N=275)

| Number of Evaluations | Percentage of Evaluators |
|---|---|
| 1 | 65% |
| 2 | 15 |
| 3 | 7 |
| 4-6 | 8 |
| 7-9 | 2 |
| 10 or more | 3 |
| Total | 100% |

Of those 275 evaluators, 249 completed at least one Project Evaluator Questionnaire. We looked to see if the characteristics of evaluators would look different if the evaluator was used as the unit of analysis. Among the findings were that:

- 50 percent of evaluators had a Ph.D., 30 percent had a master's degree, 17 percent had an Ed.D., 2 percent had a bachelor's degree, and 2 percent had some other degree as their highest level of education;

- The mean number of college or university-sponsored courses in the following areas were: (a) methods for evaluating educational programs = 3.2; (b) measurement of educational progress = 3.0; (c) data analysis of educational data = 3.7; and (d) approaches to educating limited English proficient students = 4.3;

- The mean number of years of experience in evaluating educational programs was 14.3, and the mean number of years of experience in evaluating Title VII projects was 8.9.

Generally, the results using the evaluator as the unit of analysis were very similar to those using project as the unit of analysis. The fact that the results were similar indicates that evaluators who did a large number of Title VII evaluations came from backgrounds similar to those who did few such evaluations, but the former did have slightly more relevant coursework and evaluation experience.

D. Funding of Evaluators

We asked project directors to indicate how much money had been spent either for an external evaluator or for "external" evaluation by district staff. The results are presented in Table 13. Of those who reported spending any money for evaluation services, the median amount spent was $3,500, with a range from $175 to $25,000. Among those with evaluation budgets, the average project spent 3.2 percent of its overall budget for evaluation. Approximately 19 percent of project directors failed to answer this question. For most of these projects, we assume that no project funds were specifically allocated for evaluation (e.g., the evaluation was performed by project staff as part of their regular responsibilities). In fact, in only 27 percent of those cases in which no budget was provided did an external evaluator have primary responsibility for the evaluation.

TABLE 13

Amount of Funds Spent for Evaluation Services
(N=491)

| Funds for Evaluation | Percentage |
|---|---|
| Less than $1,000 | 2% |
| $1,000 to $1,999 | 10 |
| $2,000 to $2,999 | 16 |
| $3,000 to $3,999 | 26 |
| $4,000 to $4,999 | 15 |
| $5,000 to $5,999 | 16 |
| More than $6,000 | 15 |
| Total | 100% |
| Mean Funds for Evaluation | $4,022 |
| Median Funds for Evaluation | $3,500 |

We compared the amounts and proportions of project budgets spent on evaluation across different types of projects. There were no major differences based on the year of initial funding. There were differences, however, based on the size and type of project. As shown in Table 14, larger projects spent more on evaluation than did smaller projects, but they spent a smaller proportion of their overall budget. In addition, transitional projects spent more than special alternative projects, though this difference was due to their larger size.

TABLE 14

Mean Funds for Evaluation By Project Type

| Project Type | Mean Funds | Percentage of Total Budget | N |
|---|---|---|---|
| Small (less than $100,000) | $3,403 | 4.5% | 156 |
| Medium ($100,000-$149,999) | $3,614 | 2.9% | 138 |
| Large ($150,000-$199,999) | $4,433 | 2.5% | 110 |
| Very large ($200,000 or more) | $5,206 | 2.0% | 87 |
| Transitional | $4,191 | 3.1% | 385 |
| Special alternative | $3,407 | 3.6% | 106 |

# IV. MEASUREMENT OF STUDENT OUTCOMES

## A. Standardized Achievement Tests

There are a variety of ways to examine the effects of an educational project on students. One of the most popular approaches is the use of standardized achievement tests. The advantages of standardized tests are that they assess a known and consistent set of skills, that they include national norms against which to compare students at a local site, and that they have known psychometric properties (e.g., internal consistency, test-retest reliability). Their disadvantages are that they may contain material unrelated to the curriculum of a particular program, that they may be culturally or socially biased, and that the national norms may not be relevant to the project group under study.

We asked project evaluators to describe if and how standardized achievement tests were used in Title VII project evaluations. According to evaluators (N=553), 93 percent of evaluations used norm-referenced standardized achievement tests. We compared different types of projects to determine if there were differences in the use of such tests. We found no major differences based on the type of project (transitional vs. special alternative), year of initial funding, or size of grant.

We also asked in what achievement areas standardized tests were used. As shown in Table 15, the achievement areas most often tested were English reading, English language arts, and mathematics. A typical evaluation included achievement testing in three or four of the areas listed in Table 15.

TABLE 15

Areas of Standardized Achievement Testing in Title VII Evaluations
(N=553)

| Achievement Area | Percentage of Evaluations |
|---|---|
| English reading | 84% |
| English language arts | 80 |
| Mathematics | 72 |
| English writing | 34 |
| Native language arts | 28 |
| Science | 23 |
| Social studies | 22 |
| Other area | 7 |

B. Criterion Referenced Achievement Tests

A second form of testing that is often used in evaluating educational programs is criterion referenced testing. In such testing, students are compared against pre-established standards of achievement rather than against national norms. Criterion referenced testing is often related to specific curricular objectives, and, as such, it may be a better assessment of classroom achievement than standardized tests. To the extent that criterion referenced tests are closely related to curricular content, they may also be less likely to include cultural or social bias. A disadvantage of criterion referenced tests is that without a valid and reliable comparison group it is impossible to judge the educational significance of achievement gains (e.g., a group having no special treatment may show similar gains). Also, if the test employed is not commonly used, outsiders may have difficulty in assessing the educational worth of any gains that are achieved.

We asked project evaluators if and how criterion referenced tests were used in Title VII evaluations. Of those responding (N=551), 39 percent of evaluations were reported to use criterion referenced achievement testing. We made comparisons based on project type, year of initial funding, and size of grant and found no major differences in the use of criterion referenced tests.

For those projects that used criterion referenced tests, we also asked in what achievement areas they were used. The results are presented in Table 16. There is a much more even distribution across achievement areas than for standardized achievement tests, though no one area was tested in as many as one-quarter of the evaluations.

TABLE 16

Areas of Criterion Referenced Achievement Testing in Title VII Evaluations
(N=551)

| Achievement Area | Percentage of Evaluations |
|---|---|
| English reading | 22% |
| English writing | 21 |
| English language arts | 19 |
| Mathematics | 17 |
| Native language arts | 11 |
| Social studies | 11 |
| Science | 10 |
| Cultural awareness | 8 |
| Other area | 3 |

C. English Oral Proficiency Tests

An important objective of many Title VII projects is to improve the English oral proficiency of project students. Therefore, we asked project evaluators to indicate whether English oral proficiency tests were used as parts of Title VII evaluations. Of those responding (N=546), 72 percent reported that such tests were used in the evaluation. When we compared responses on this item across different types of projects based on program type, year of initial funding, and size of grant, the only significant difference we found was for the largest projects ($200,000 or more)(N=111), which used English oral proficiency tests in only 60 percent of cases. We believe that for large projects (which serve large numbers of students), the cost of individually-administered oral proficiency tests may have been a barrier.

We also asked evaluators to describe the oral proficiency areas that were tested. Table 17 shows the results on this question. Aural comprehension, fluency, and vocabulary were all frequently tested. The "other" category in Table 17 included pronounciation (2 percent of evaluations), syntax (2 percent) and complexity of language (1 percent).

TABLE 17

Areas of English Oral Proficiency Testing in Title VII Evaluations
(N=546)

| Oral Proficiency Area | Percentage of Evaluations |
|---|---|
| Aural comprehension | 60% |
| Fluency | 60 |
| Vocabulary | 54 |
| Other area | 5 |

D. Other Measures of Student Academic Achievement

In addition to standardized test data, there are a variety of other measures which can be used to assess the effectiveness of a project in improving student academic achievement. For example, one of the major current initiatives in the field of educational assessment is the development of valid and reliable methods for rating student products such as portfolios.

We asked evaluators to indicate what measures of student academic achievement other than standardized tests were used in Title VII evaluations. The responses are presented in Table 18. None of these measures were used as frequently as standardized tests, though classroom grades were used in almost a third of the evaluations. Among the "other" measures, the most frequently mentioned was writing samples (1 percent of evaluations).

TABLE 18

Other Measures of Student Academic Achievement Used in Title VII Evaluations
(N=546)

| Measure of Achievement | Percentage of Evaluations |
|---|---|
| Classroom grades | 32% |
| Teacher developed tests | 19 |
| Curriculum-based tests | 18 |
| Evaluation of student products (e.g., portfolios) | 15 |
| Credit accrual | 8 |
| Other measure | 4 |

E. Other Student Outcomes

Title VII projects have a variety of objectives for students in addition to improving academic achievement. Many of these objectives are clearly related to academic achievement, however. For example, a project may want to improve students' school attendance, under the assumption that attendance is a necessary precursor to achievement.

We asked evaluators to indicate what other types of student outcomes were assessed as parts of Title VII evaluations. As the results in Table 19 indicate, the other student outcomes most frequently measured were attendance, retention/promotion, and exit from the Title VII program. The most frequently mentioned "other" measures were time on task (3 percent of evaluations) and transition to English proficient status (1 percent).

TABLE 19

Other Student Outcomes Measured in Title VII Evaluations
(N=560)

| Measure of Student Outcome | Percentage of Evaluations |
|---|---|
| Absenteeism/attendance | 63% |
| Grade retention/promotion | 55 |
| Exit from the Title VII program | 49 |
| Referral to special education | 46 |
| Placement in gifted and talented programs | 42 |
| School dropout | 42 |
| Student self-esteem | 27 |
| Attitudes toward school | 24 |
| Cultural pride | 21 |
| High school graduation | 21 |
| Enrollment in postsecondary schools | 14 |
| Exit from other special programs | 7 |
| Other measure | 10 |

F. Students Included in Evaluation

In addition to the question of what student data are collected in an evaluation, equally important are the questions of from whom the data are collected and what comparisons are made using the data. We asked evaluators a series of questions about comparison groups and comparative analyses relating to student outcomes.

For each of the types of student achievement data described in the preceding sections of this chapter, we asked evaluators to indicate about whom such data were collected in the evaluation. The results are presented in Table 20. The first column in the table shows the percentages of all Title VII evaluations in which specific categories of students were included in data collection. The second column in the table shows the percentages of evaluations using a particular measure in which specific student groups were included (i.e., if an evaluation did not include norm referenced achievement tests for any students, it is not included in the second column for that measure).

As the table shows, local comparison groups were used for any purpose in only 34 percent of the evaluations. Also, although required by federal regulation, data were collected concerning former project participants in only 28 percent of evaluations. We compared projects based on year of initial funding, project type, and project size, and found no major differences on these variables.

TABLE 20

Types of Students Included in Analyses of Student Achievement

| Achievement Measure/ Student Type | Percentage of Evaluations | Percentage of Evaluations Using Measure* |
|---|---|---|
| Norm referenced achievement tests (N=553) | | |
| LEP students in the project | 88% | 95% |
| EP students in the project | 32 | 35 |
| LEP students in a comparison group | 18 | 20 |
| All students in the district | 33 | 36 |
| Former LEP students no longer in the project | 24 | 26 |
| Criterion referenced achievement tests (N=551) | | |
| LEP students in the project | 38% | 97% |
| EP students in the project | 15 | 39 |
| LEP students in a comparison group | 8 | 20 |
| All students in the district | 12 | 32 |
| Former LEP students no longer in the project | 7 | 18 |
| English oral proficiency tests (N=546) | | |
| LEP students in the project | 71% | 99% |
| LEP students in a comparison group | 12 | 17 |
| Former LEP students no longer in the project | 11 | 15 |
| Other measures of academic achievment (N=546) | | |
| LEP students in the project | 49% | 92% |
| EP students in the project | 16 | 30 |
| LEP students in a comparison group | 6 | 12 |
| All students in the district | 11 | 21 |
| Former LEP students no longer in the project | 8 | 16 |
| All measures of academic achievement (N=560) | | |
| LEP students in the project | 94% | 97% |
| EP students in the project | 34 | 35 |
| LEP students in a comparison group | 23 | 24 |
| All students in the district | 39 | 40 |
| Former LEP students no longer in the project | 28 | 29 |

* Includes only those evaluations in which the achievement measure is used with any students.

G. Analytic Comparisons

We also asked project evaluators to describe the types of analytic comparisons that had been made with student outcome data. For example, we asked if pre-post comparisons of project students had been performed for each of the types of data that had been collected. The results on these questions are presented in Table 21. As in Table 20, the first column shows the overall percentages of evaluations making a specific comparison, while the second column shows the percentages of evaluations using the measure for any students which made the comparison.

The results indicate that pre-post comparisons of project students were the comparisons most frequently made. For norm referenced achievement tests, comparisons with national or state norms were also applied in two-thirds of the evaluations. Gap reduction analysis was used in about one-third of the evaluations.

We also asked project directors when various types of testing were done for the evaluation (fall only, winter/spring only, both times, neither time). It should be noted that the Title VII evaluation regulations call for administration of the evaluation instruments " at twelve-month testing intervals." These results are presented in Table 22. Generally, standardized achievement tests were most likely to be given in the winter/spring, English oral proficiency tests and criterion based achievement tests throughout the year, and other student measures (self-concept, school attitudes, etc.) in the fall only.

TABLE 21

Types of Analytic Comparisons Performed Using Student Outcome Measures

| Outcome Measure/ Analytic Comparison | Percentage of Evaluations | Percentage of Evaluations Using Measure* |
|---|---|---|
| Norm referenced achievement tests (N=553) | | |
| Pre-post comparisons of project students | 78% | 84% |
| Comparisons with national or state norms | 62 | 67 |
| Comparisons with non-project LEP students in the district | 18 | 20 |
| Comparisons with overall school or district norms | 25 | 27 |
| Gap reduction analysis | 34 | 37 |
| Other | 4 | 5 |
| Criterion referenced achievement tests (N=551) | | |
| Pre-post comparisons of project students | 33% | 85% |
| Comparisons with national or state norms | 11 | 27 |
| Comparisons with non-project LEP students in the district | 9 | 22 |
| Comparisons with overall school or district norms | 11 | 27 |
| Gap reduction analysis | 7 | 19 |
| Other | 2 | 5 |
| English oral proficiency tests (N=546) | | |
| Pre-post comparisons of project students | 61% | 85% |
| Comparisons with non-project LEP students in the district | 12 | 16 |
| Other | 3 | 4 |
| Other measures of academic achievement (N=546) | | |
| Pre-post comparisons of project students | 35% | 66% |
| Comparisons with non-project LEP students in the district | 6 | 12 |
| Comparisons with overall school or district results | 11 | 20 |
| Other | 4 | 7 |
| Other measures of student outcomes (N=560) | | |
| Pre-post comparisons of project students | 51% | 59% |
| Comparisons with non-project LEP students in the district | 13 | 15 |
| Comparisons with overall school or district results | 32 | 38 |
| Other | 6 | 7 |

TABLE 21

(continued)

| Outcome Measure/ Analytic Comparison | Percentage of Evaluations | Percentage of Evaluations Using Measure* |
|---|---|---|
| All measures of student outcomes (N=560) | | |
| Pre-post comparisons of project students | 90% | 93% |
| Comparisons with national or state norms | 62 | 64 |
| Comparisons with non-project LEP students in the district | 28 | 29 |
| Comparisons with overall school or district norms | 45 | 47 |
| Gap reduction analysis | 35 | 36 |
| Other | 13 | 13 |

* Includes only those evaluations in which the achievement measure is used with any students.

TABLE 22

Timing of Student Testing for Various Measures of Student Outcomes
(N=608)

| | Type of Testing | | | |
|---|---|---|---|---|
| When Students Were Tested | Standardized Achievement Tests | Criterion-Based Tests | English Oral Proficiency Tests | Other Student Measures (e.g., self-concept) |
| Fall only | 5% | 5% | 14% | 3% |
| Winter/spring only | 58 | 21 | 13 | 18 |
| Both fall and winter/spring | 37 | 46 | 63 | 35 |
| Neither fall nor winter/spring | 6 | 28 | 10 | 44 |
| Total | 100% | 100% | 100% | 100% |

# V. EVALUATION IMPLEMENTATION

## A. Evaluation Planning

Evaluations of Title VII projects can involve a broad range of persons and activities. Careful planning of the evaluation thus is extremely important for producing a quality process and product. Therefore, we asked project directors to describe the persons who were involved in such planning.

First, we asked for a listing of the persons who were significantly involved in developing the evaluation plan in the original application for the project. Table 23 shows the results on that question. "Outside" evaluators were the persons most likely to be involved, although project directors were also significantly involved in more than half of the projects.

TABLE 23

Persons Involved in Developing Evaluation Plans for Original Title VII Applications (N=603)

| Type of Person | Percentage of Projects |
|---|---|
| Present project director | 58% |
| Former project director(s) | 36 |
| Other project staff member(s) | 39 |
| District evaluation/testing staff | 22 |
| Other district staff | 27 |
| "Outside" evaluator or technical assistance provider | 65 |
| Other outside expert | 10 |

We also asked if there had been any major revisions or changes to the original evaluation plan. Of those responding (N=598), only 25 percent reported that major revisions had been made. For those projects which had made revisions, we asked who was involved in making the revisions. The results are presented in Table 24. The percentages in Table 24 are of those projects in which revisions had been made. In general, the same types of persons were involved in making revisions as were involved in original evaluation plans although present project directors were more likely to be involved in revisions.

TABLE 24

Persons Involved in Making Revisions to Original Evaluation Plans
(N=165)

| Type of Person | Percentage of Projects |
|---|---|
| Present project director | 72% |
| Former project director(s) | 30 |
| Other project staff member(s) | 45 |
| District evaluation/testing staff | 28 |
| Other district staff | 22 |
| "Outside" evaluator or technical assistance provider | 69 |
| Other outside expert | 14 |

B. Responsibilities for Evaluation Activities

On the survey of Title VII project directors, we asked a very detailed question about the involvement of various categories of persons in specific evaluation activities. The question listed five categories of persons (project director, other project staff, district testing or evaluation staff, other district staff, and outside evaluator) and 21 specific activities, and requested the level of involvement by each category of person in each activity. Respondents were asked to use a three-point rating scale, with "1" indicating primary responsibilty, "2" indicating some involvement, and "3" indicating no involvement. Respondents were asked to use the "1" rating (primary responsibility) for only one category of person.

The results on this question can be presented in a number of ways. One complicating factor is that some respondents ignored our request and used the "1" rating for more than one category of person. Thus, their responses are not directly comparable to those of the respondents who followed our direction.

The results on this question are presented in Tables 25, 26, and 27. Table 25 shows the percentage of projects where the person(s) had any involvement in the activity (i.e., there was a "1" or "2" rating). Table 26 shows the percentage of projects in which the person(s) had primary responsibility (i.e., a "1" rating). It should be noted that the percentages in Table 26 may include more than one category of person per activity or may include no one (e.g., if the activity was not performed in that year). Finally, in Table 27 we describe the mean involvement rating for each type of person in each activity. In order to make the projects more comparable, for Table 27, we have recoded those "1" ratings, which were given to more than one category of person as "1.5" ratings (i.e., halfway between "primary responsibilty" and "some involvement").

TABLE 25

Involvement in Evaluation Activities by Project and District Staff-
Percentage of Projects in Which Staff Have Some Involvement
(N=603)

| Evaluation Activity | Category of Person | | | | |
|---|---|---|---|---|---|
| | Project Director | Other Project Staff | District Testing Staff | Other District Staff | Outside Evaluator |
| a. Selecting the evaluation questions | 71% | 42% | 21% | 18% | 72% |
| b. Selecting the evaluation design | 70 | 34 | 22 | 16 | 73 |
| c. Selecting tests and other instruments to be used | 80 | 46 | 36 | 26 | 54 |
| d. Developing new instruments to be used in the evaluation | 52 | 32 | 18 | 16 | 51 |
| e. Testing students using standardized achievement tests | 55 | 63 | 41 | 47 | 20 |
| f. Testing students on oral proficiency | 52 | 73 | 22 | 43 | 15 |
| g. Testing students on other measures (criterion-referenced tests, etc.) | 42 | 64 | 24 | 46 | 15 |
| h. Collecting student data from existing records (grades, dropouts, referrals, etc.) | 73 | 74 | 26 | 38 | 34 |
| i. Observing and evaluating project-related classroom activities | 87 | 60 | 19 | 34 | 59 |
| j. Observing and evaluating other project-related student activities | 84 | 67 | 19 | 37 | 52 |
| k. Observing and evaluating activities for teachers | 86 | 56 | 16 | 39 | 52 |
| l. Observing and evaluating activities for parents | 86 | 70 | 13 | 34 | 46 |
| m. Collecting records (class lists, attendance logs, rating sheets) from teacher activities | 76 | 74 | 18 | 32 | 34 |
| n. Collecting records (attendance logs, rating sheets) from parent activities | 76 | 76 | 11 | 29 | 32 |
| o. Collecting and assessing project-developed curricular materials | 75 | 67 | 11 | 29 | 40 |
| p. Summarizing evaluation data (developing tables, etc.) | 53 | 32 | 33 | 10 | 80 |
| q. Performing statistical analyses of data | 33 | 15 | 23 | 8 | 82 |
| r. Drafting evaluation report(s) | 48 | 22 | 19 | 6 | 82 |
| s. Preparing evaluation conclusions and recommendations | 51 | 23 | 18 | 7 | 82 |
| t. Presenting evaluation results and conclusions to project and school staff | 80 | 37 | 18 | 10 | 62 |
| u. Presenting evaluation results and conclusions to school administrators | 86 | 33 | 17 | 10 | 52 |

TABLE 26

Involvement in Evaluation Activities by Project and District Staff-
Percentage of Projects in Which Staff Have Primary Responsibility
(N=603)

| Evaluation Activity | Category of Person | | | | |
|---|---|---|---|---|---|
| | Project Director | Other Project Staff | District Testing Staff | Other District Staff | Outside Evaluator |
| a. Selecting the evaluation questions | 34% | 5% | 9% | 2% | 55% |
| b. Selecting the evaluation design | 32 | 4 | 10 | 4 | 58 |
| c. Selecting tests and other instruments to be used | 50 | 10 | 21 | 7 | 26 |
| d. Developing new instruments to be used in the evaluation | 24 | 7 | 8 | 2 | 39 |
| e. Testing students using standardized achievement tests | 25 | 34 | 27 | 23 | 7 |
| f. Testing students on oral proficiency | 23 | 51 | 11 | 22 | 4 |
| g. Testing students on other measures (criterion-referenced tests, etc.) | 18 | 40 | 14 | 25 | 4 |
| h. Collecting student data from existing records (grades, dropouts, referrals, etc.) | 43 | 45 | 10 | 8 | 12 |
| i. Observing and evaluating project-related classroom activities | 53 | 27 | 7 | 7 | 32 |
| j. Observing and evaluating other project-related student activities | 52 | 34 | 5 | 9 | 20 |
| k. Observing and evaluating activities for teachers | 63 | 15 | 5 | 12 | 19 |
| l. Observing and evaluating activities for parents | 63 | 34 | 4 | 7 | 13 |
| m. Collecting records (class lists, attendance logs, rating sheets) from teacher activities | 50 | 45 | 6 | 9 | 9 |
| n. Collecting records (attendance logs, rating sheets) from parent activities | 51 | 45 | 3 | 9 | 8 |
| o. Collecting and assessing project-developed curricular materials | 53 | 34 | 2 | 6 | 15 |
| p. Summarizing evaluation data (developing tables, etc.) | 18 | 5 | 13 | 2 | 75 |
| q. Performing statistical analyses of data | 9 | 2 | 14 | 2 | 80 |
| r. Drafting evaluation report(s) | 14 | 2 | 12 | 1 | 79 |
| s. Preparing evaluation conclusions and recommendations | 12 | 2 | 12 | 1 | 80 |
| t. Presenting evaluation results and conclusions to project and school staff | 51 | 9 | 10 | 1 | 45 |
| u. Presenting evaluation results and conclusions to school administrators | 65 | 8 | 10 | 2 | 29 |

TABLE 27

Involvement in Evaluation Activities by Project and District Staff-
Mean Level of Involvement
(1 = primary responsibilty, 1.5 = shared primary responsibility,
2 = some involvement, 3 = no involvement)
(N=603)

| Evaluation Activity | Category of Person | | | | |
|---|---|---|---|---|---|
| | Project Director | Other Project Staff | District Testing Staff | Other District Staff | Outside Evaluator |
| a. Selecting the evaluation questions | 2.0 | 2.6 | 2.7 | 2.8 | 1.8 |
| b. Selecting the evaluation design | 2.0 | 2.6 | 2.7 | 2.8 | 1.7 |
| c. Selecting tests and other instruments to be used | 1.8 | 2.5 | 2.5 | 2.7 | 2.2 |
| d. Developing new instruments to be used in the evaluation | 2.3 | 2.6 | 2.8 | 2.8 | 2.1 |
| e. Testing students using standardized achievement tests | 2.2 | 2.1 | 2.4 | 2.3 | 2.8 |
| f. Testing students on oral proficiency | 2.3 | 1.8 | 2.7 | 2.4 | 2.8 |
| g. Testing students on other measures (criterion-referenced tests, etc.) | 2.4 | 2.0 | 2.6 | 2.3 | 2.8 |
| h. Collecting student data from existing records (grades, dropouts, referrals, etc.) | 1.9 | 1.9 | 2.6 | 2.6 | 2.6 |
| i. Observing and evaluating project-related classroom activities | 1.7 | 2.2 | 2.8 | 2.6 | 2.2 |
| j. Observing and evaluating other project-related student activities | 1.7 | 2.1 | 2.8 | 2.6 | 2.3 |
| k. Observing and evaluating activities for teachers | 1.6 | 2.2 | 2.8 | 2.5 | 2.4 |
| l. Observing and evaluating activities for parents | 1.6 | 2.0 | 2.8 | 2.6 | 2.4 |
| m. Collecting records (class lists, attendance logs, rating sheets) from teacher activities | 1.8 | 1.9 | 2.8 | 2.6 | 2.6 |
| n. Collecting records (attendance logs, rating sheets) from parent activities | 1.8 | 1.9 | 2.9 | 2.6 | 2.6 |
| o. Collecting and assessing project-developed curricular materials | 1.8 | 2.0 | 2.9 | 2.7 | 2.5 |
| p. Summarizing evaluation data (developing tables, etc.) | 2.3 | 2.6 | 2.6 | 2.9 | 1.5 |
| q. Performing statistical analyses of data | 2.6 | 2.8 | 2.6 | 2.9 | 1.4 |
| r. Drafting evaluation report(s) | 2.4 | 2.8 | 2.7 | 2.9 | 1.4 |
| s. Preparing evaluation conclusions and recommendations | 2.4 | 2.8 | 2.7 | 2.9 | 1.4 |
| t. Presenting evaluation results and conclusions to project and school staff | 1.8 | 2.6 | 2.7 | 2.9 | 2.0 |
| u. Presenting evaluation results and conclusions to school administrators | 1.6 | 2.6 | 2.8 | 2.9 | 2.2 |

The result in these three tables suggest that project directors were most involved in process evaluation activities (observing and evaluating classroom, teacher, and parent activities), but outside evaluators were most involved in outcome evaluation (summarizing evaluation data, performing statistical analyses, and drafting reports). Other project staff were heavily involved in testing students, but other district staff were moderately involved in student testing. Project directors and outside evaluators shared the responsibility for presenting results to project and school staff and school administrators, but project directors had a more active role with school administrators.

C. Time Expended on Evaluation Activities

We asked project directors about how much project staff time had been devoted to specific evaluation activities. The time to be reported was not to include time by nonproject staff or outside evaluators and was to be reported in person/days (a person/day was defined as the equivalent of one individual working 7 or 8 hours). The responses to this item suggested that a number of respondents had trouble understanding what we were asking for, so we had to edit a large number of responses and delete some others. Thus, the data from this question may be considerably less precise than for other questions.

The results are presented in Table 28. As can be seen, the largest amount of staff time was devoted to student testing, followed by collection of existing student records.

TABLE 28

Project Staff Time Spent on Evaluation Activities

| Activity | N | Median Days | Mean Days |
|---|---|---|---|
| Evaluation planning | 549 | 4 | 8.0 |
| Student testing | 548 | 12 | 24.9 |
| Collection of existing student records | 541 | 8 | 15.4 |
| Collection of other evaluation data (observation, assessment of parent sessions) | 541 | 7 | 13.6 |
| Analyzing data and writing reports | 537 | 5 | 10.0 |
| Reviewing reports and revising project plans | 541 | 5 | 7.6 |

We added the amounts of staff time devoted to specific evaluation activities together to get an overall measure. The project mean was 79.7 person/days (N=528). We also compared this total across projects based on year of initial funding, type of project, and project size. The only significant difference was that larger projects devoted more staff time to evaluation activities (less than $100,000: mean = 68.1 days, N=173; $100,000-149,999: mean = 75.9 days, N=139; $150,000-199,999: mean = 86.1 days, N=117; $200,000 or more: mean = 97.7 days, N=99).

We also asked evaluators how many hours they (or their evaluation staff) spent on various evaluation activities for the project. The results are presented in Table 29. Evaluators spent most of their time analyzing data and preparing reports.

TABLE 29

Evaluator Time Spent on Evaluation Activities

| Activity | N | Median Hours | Mean Hours |
|---|---|---|---|
| Collecting data from school records | 517 | 16 | 26.4 |
| Observing project activities | 523 | 16 | 25.2 |
| Interviewing project and school staff | 538 | 10 | 16.7 |
| Analyzing data and preparing reports | 547 | 60 | 79.5 |

We added the hours from Table 29 together to get an overall measure. The project mean was 144.1 hours. We also compared these totals across types of projects, and found that evaluators devoted more time to larger projects (less than $100,000: mean = 129.5 hours, N=156; $100,000-149,999: mean = 138.0 hours, N=135; $150,000-199,999: mean = 156.0 hours, N=109; $200,000 or more: mean = 163.2 hours, N=94). Also, evaluators spent more time on transitional projects (mean = 150.1, N=388) than on special alternative projects (mean = 105.6, N=106), though this difference may have been at least partially a function of project size.

Finally, we asked evaluators how many hours they (or their evaluation staff) had spent on-site related to the evaluation of the project. The median number of hours was 35 (N=535), and the mean was 58.6. In 24 percent of projects, the evaluation staff had spent 16 or fewer hours on-site. There were significant differences on this question across types of projects. As Table 30 shows, projects in their last year of funding (those initially funded in 1985), large projects, and transitional projects all had higher than average mean amounts of evaluator time on site.

TABLE 30

Amount of Evaluator Time Spent On Site By Category of Project

| Category of Project | N | Mean Hours |
|---|---|---|
| Initial funding - 1985 | 112 | 73.7 |
| Initial funding - 1986 | 82 | 48.8 |
| Initial funding - 1987 | 54 | 53.2 |
| Initial funding - 1988 | 111 | 61.3 |
| Initial funding - 1989 | 176 | 53.4 |
| Transitional | 417 | 62.4 |
| Special alternative | 118 | 45.2 |
| Less than $100,000 | 161 | 52.3 |
| $100,000-149,999 | 144 | 56.8 |
| $150,000-199,999 | 122 | 51.9 |
| $200,000 or more | 108 | 77.8 |

# VI. USE OF OUTSIDE RESOURCES

## A. Use of the Bilingual Education Evaluation System

Title VII projects have been provided a number of resources by the U.S. Department of Education to assist in their evaluations. One of these resources is the Bilingual Education Evaluation System (BEES) which was developed by RMC Research Corporation. This system includes a summary of federal regulations, a User's Guide for conducting evaluations, and a technical manual.

We asked project directors and evaluators a series of questions concerning their knowledge and use of BEES. Table 31 shows the responses of both groups on a question concerning knowledge of BEES. Only approximately one-quarter of project directors had read BEES, while more than two-third of evaluators had done so.

TABLE 31

Knowledge of the Bilingual Education Evaluation System By Title VII Project Directors and Evaluators

| Level of Knowledge | Project Directors (N=601) | Evaluators (N=553) |
|---|---|---|
| I am not aware of this document. | 61% | 24% |
| I am aware of the document, but have not read it. | 11 | 7 |
| I have looked through the document but have not read it thoroughly. | 17 | 15 |
| I have read the document thoroughly. | 10 | 55 |
| Total | 100% | 100% |

We next asked about the extent to which BEES had been used in designing or conducting the evaluation of the project. The results are presented in Table 32. Those project directors and evaluators who had not heard of BEES were told to skip this item. As might be expected based on responses to the previous item, evaluators were much more likely to report that BEES had been used in the evaluation than were project directors.

TABLE 32

Use of BEES in Designing or Conducting Title VII Evaluations

| Extent of Use | Project Directors (N=608) | Evaluators (N=560) |
|---|---|---|
| The BEES had a major impact on the evaluation. | 9% | 24% |
| A few concepts from BEES were used in the evaluation. | 14 | 34 |
| The BEES was not used in the evaluation. | 11 | 19 |
| I don't know/not applicable/missing. | 67 | 23 |
| Total | 100% | 100% |

For those project directors and evaluators who indicated knowledge of BEES, we also asked about the usefulness of BEES as it related to their project. The results are presented in Table 33. BEES was rated as "moderately useful" or "very useful" by more than 80 percent of the 122 project directors and 343 evaluators who were willing to rate it.

For evaluators only, we asked what concepts or methods from the BEES they had found to be particularly useful. The most frequently mentioned open-ended responses were: gap reduction (33 percent of all responses); report formats (17 percent); evaluation design/comparison groups (12 percent); Title VII requirements (11 percent); data collection approaches (5 percent); data analysis (4 percent); and process evaluation (4 percent). Of all respondents to the questionnaire, 44 percent gave at least one response to this item.

TABLE 33

Usefulness of BEES in Designing or Conducting Title VII Evaluations

| Usefulness | Project Directors (N=608) | Evaluators (N=560) |
|---|---|---|
| It is not very useful. | 3% | 11% |
| It is moderately useful. | 10 | 32 |
| It is very useful. | 8 | 18 |
| I don't know/not applicable/missing. | 80 | 39 |
| Total | 100% | 100% |

B. Use of the Evaluation Assistance Centers

A second resource which has been provided to Title VII grantees to assist in evaluation is the presence of the two Evaluation Assistance Centers (EACs). These centers are funded by the U.S. Department of Education. They are responsible for providing training and technical assistance in evaluation design and evaluation methods to Title VII projects. As part of the survey, therefore, we asked project directors and evaluators about their experiences with the EACs.

We asked the project directors if their projects had received any assistance from the EACs. Of those responding (N=600), 65 percent said that they had received some assistance. There were no major differences in this variable based on the year of initial funding, the type of grant, or the size of grant. We also asked evaluators if they had received such assistance. Of those responding (N=555), 65 percent of them also reported that they had received some assistance. When the responses of project directors and evaluators were combined (N=608), the results indicated that 82 percent of all projects had someone who had received assistance from an EAC.

We asked project directors who specifically had received the assistance from the EAC. The project directors reported that the assistance was most frequently received by project directors (59 percent of all projects), followed by other project-supported staff (24 percent), "outside" evaluators (22 percent), district testing and evaluation staff (12 percent), and other district staff (9 percent).

Table 34 shows the nature of the assistance received. As reported by project directors and evaluators, materials sent by mail and multi-project training sessions were the most frequent forms of assistance, though telephone consultation and on-site technical assistance were also often employed. (The percentages in the table are of all questionnaire respondents.)

TABLE 34

Types of Assistance Received From Evaluation Assistance Centers

| Type of Assistance | Project Directors (N=608) | Evaluators (N=560) |
|---|---|---|
| Materials sent by mail | 43% | 41% |
| Training sessions including staff from more than one project | 41 | 42 |
| Telephone consultation | 34 | 37 |
| On-site training or assistance specific to the project | 19 | 13 |
| Other | 7 | 8 |

For evaluators only, we asked about the content areas of the assistance. As Table 35 shows, the most frequently received content area was evaluation design, including comparison groups and gap reduction. (The perercentages in this table are also for all questionnaire respondents.)

We asked both project directors and evaluators to indicate when their most recent contact was with the EAC. As the results in Table 36 show, most had had contact with the EAC since the beginning of 1990. For project directors, the median most recent contact was in October of 1990, but for evaluators, the median was in January of 1991.

Finally, we asked project directors and evaluators to rate the usefulness of the assistance that they had received from the EAC. The results are presented in Table 37. Approximately equal proportions of respondents rated the assistance as "moderately useful" and "very useful," and very few rated it as "not very useful."

We compared the ratings of usefulness across different categories of projects and found no major differences in ratings based on year of initial funding, type of project, or size of project.

TABLE 35

Content Areas Of Assistance Received From EACs
(N=560)

| Content Area of Assistance | Percentage of Projects |
|---|---|
| Evaluation design (comparison groups, gap reduction, etc.) | 52% |
| Methods for recording evaluation data (summary sheets, computer files, etc.) | 31 |
| Analysis of evaluation data (statistical techniques, selection of analytic subgroups) | 28 |
| Selection of existing evaluation instruments (achievement tests, criterion-referenced tests, etc.) | 25 |
| Data collection methodologies (administration of achievement tests, classroom observation, etc.) | 23 |
| Use of evaluation results (feedback to staff, using results to modify program elements, etc.) | 20 |
| Development of new evaluation instruments (questionnaires, rating sheets, etc.) | 16 |

TABLE 36

Most Recent Contact With An EAC

| Most Recent Contact | Project Directors (N=342) | Evaluators (N=332) |
|---|---|---|
| 1987 or earlier | 3% | 3% |
| 1988 | 4 | 3 |
| 1989 | 12 | 10 |
| 1990 | 44 | 30 |
| 1991 | 37 | 53 |
| Total | 100% | 100% |

TABLE 37

Usefulness of the Assistance Received From EACs

| Usefulness of Assistance | Project Directors (N=380) | Evaluators (N=357) |
|---|---|---|
| Not very useful | 4% | 11% |
| Moderately useful | 43 | 41 |
| Very useful | 53 | 48 |
| Total | 100% | 100% |

C. Use of Other Materials and Resources

In addition to the Bilingual Education Evaluation System and the Evaluation Assistance Centers, we were also interested in other resources which Title VII projects used to design and implement their evaluations. Therefore, we asked about materials and sources of technical assistance that had been used.

We asked both project directors and evaluators about other evaluation-related materials. In describing such materials, we gave as an example the Title I Evaluation Reporting System (TIERS). Table 38 shows the most frequent open-ended responses which were provided.

TABLE 38

Evaluation-Related Materials Used By Title VII Projects

| Materials Listed | Percentage of All Respondents | |
|---|---|---|
| | Project Directors (N=608) | Evaluators (N=560) |
| Title I Evaluation Reporting System (TIERS) | 1% | 11% |
| EAC Evaluation Report Checklist | 1 | 3 |
| Program Evaluation Kit (Sage) | 0 | 3 |
| Federal regulations | 1 | 2 |
| State Department of Education materials | 1 | 2 |
| Interaction Model of Bilingual Education (Cummins) | 0 | 1 |
| RMC early publications | 0 | 1 |

We also asked project directors and evaluators if they had received any assistance specifically related to evaluating the project from sources other than the EACs. For both project directors (N=582) and evaluators (N=536), 40 percent of respondents indicated that they had received such assistance. Table 39 shows the most frequently reported sources of such assistance. Outside evaluators and consultants and state departments of eduction are the two sources most frequently used.

TABLE 39

Sources of Evaluation-Related Technical Assistance Used By Title VII Projects

| Source of Assistance | Percentage of All Respondents | |
|---|---|---|
| | Project Directors (N=608) | Evaluators (N=560) |
| Outside evaluator/consultant | 14% | 9% |
| State Department of Education | 12 | 14 |
| Multifunctional Resource Center (MRC) | 8 | 3 |
| School district evaluation staff | 7 | 7 |
| OBEMLA | 3 | 5 |
| Local university | 2 | 7 |
| Consulting firm | 1 | 2 |
| Other school district | 1 | 2 |
| County department of education | 1 | 1 |

D. Contacts With the U.S. Department of Education

Finally, we asked project directors and evaluators how often in the year they had spoken to the project monitor at the U.S. Department of Education about the evaluation of the project. The results are presented in Table 40. Most evaluators had not spoken to the project monitor, but most project directors had done so two or less times.

TABLE 40

Number of Contacts With the U.S. Department of Education Project Monitor

| Number of Contacts in Year | Project Directors (N=519) | Evaluators (N=413) |
|---|---|---|
| 0 | 29% | 75% |
| 1 | 25 | 13 |
| 2 | 22 | 8 |
| 3-5 | 18 | 3 |
| 6 or more | 6 | 1 |
| Total | 100% | 100% |

# VII. EVALUATION REPORTING

## A. Nature of Written Reports

Title VII projects are required to submit yearly evaluation reports to the U.S. Department of Education. Evaluators may also submit other special reports to the project or to the district.

We asked project directors if a written evaluation report had been prepared that described the overall results of the 1989-90 project year. Of those responding (N=604), 98 percent said that such a report had been prepared. We also asked if in addition to the overall report, any other special evaluation reports had been prepared concerning the 1989-90 project year. According to project directors (N=597), in 23 percent of projects at least one such special report had been prepared. The topics of such reports as defined in open-ended responses are shown in Table 41. Interim reports were those most frequently prepared.

TABLE 41

Types of Special Evaluation Reports Prepared

| Type of Report | Percentage of All Projects (N=608) |
|---|---|
| Interim report | 8% |
| Report to state department of education | 4 |
| Status report on LEP students in district | 3 |
| Conclusions and recommendations | 3 |
| Report on in-service activities | 3 |
| Report to local school board | 2 |
| Process/site-visit report | 2 |
| Report to parents/community | 2 |

We also asked evaluators what reports they had prepared in 1989-90 based on their evaluation activities. Their responses are shown in Table 42. Evaluators were much more likely to report the preparation of interim and process reports than were project directors. This may be due to different interpretation of what constituted an interim or process report or to the fact that project directors had to list such reports in an open-ended fashion.

TABLE 42

Types of Evaluation Reports Prepared By Evaluators

| Type of Report | Percentage of All Projects (N=560) |
|---|---|
| An end-of-year summative evaluation report | 96% |
| One or more brief feedback reports to the project director | 41 |
| One or more brief feedback reports to the project staff or project teachers | 35 |
| One or more midyear process evaluation reports | 31 |

B. Contents of Evaluation Reports

Title VII regulations require the collection and reporting of a significant amount of information concerning the project, the project staff, and project students. Therefore, we were very interested in the types and amounts of information that were collected by projects, and how much of that information was included in yearly summary reports.

We asked evaluators who was primarily responsible for determining the content and outline of evaluation reports for the year. Among those responding (N=541), 74 percent said that the evaluator had that responsibility. A total of 14 percent said that the project director had the responsibility, 2 percent said someone else (school administrator, etc.) had the responsibility, and 11 percent refused to list just one person and said the responsibility was shared.

In the survey of project directors, we listed 36 specific data elements that might have been collected and included in a summary report. We asked the project directors to indicate if the data element: (1) had been described in the 1989-90 report; (2) had been collected but not reported; or (3) had not been collected. The results are presented in Table 43.

We constructed four composite measures in order to summarize the data in Table 43: (1) the total number of elements of those listed that were in the 1989-90 report; (2) the total number of elements of those listed on which the evaluation collected information; (3) the total number of elements in the 1989-90 report from among the 24 that are specifically required by regulations; and (4) the total number of elements on which data were collected from among the 24 that are specifically required by federal regulations.

The overall results on those composites are presented in Table 44. Both in terms of all of the elements in Table 43 and those specifically related to Title VII regulations, the results indicate that the average project collected data on three-quarters of the elements and described half of the elements in their 1989-90 report.

TABLE 43

Extent To Which Specific Data Elements Are Collected and Reported in Title VII Projects

| Data Element | N | Described in 1989-90 Report | Collected But Not Reported | Not Reported | Total |
|---|---|---|---|---|---|
| Years of previous education of project students* | 590 | 39% | 27% | 34% | 100% |
| Location of previous schools of project students | 593 | 18 | 24 | 58 | 100 |
| Oral English skills of LEP students at beginning of project* | 591 | 75 | 17 | 8 | 100 |
| English reading and language skills of LEP student at beginning of project* | 587 | 80 | 13 | 7 | 100 |
| English writing skills of LEP students at beginning of project | 587 | 57 | 20 | 24 | 100 |
| Native language skills of LEP students at beginning of project | 584 | 45 | 24 | 30 | 100 |
| Academic achievement in math of LEP students at beginning of project* | 587 | 57 | 20 | 22 | 100 |
| Academic achievement in science of LEP students at beginning of project* | 584 | 26 | 18 | 56 | 100 |
| Academic achievement in social studies of LEP student at beginning of project* | 586 | 25 | 19 | 57 | 100 |
| Home language use of LEP students | 590 | 57 | 32 | 11 | 100 |
| Socioeconomic status of LEP students | 589 | 51 | 26 | 23 | 100 |
| Oral English skills of LEP students at end of project year* | 590 | 75 | 13 | 12 | 100 |
| English reading and language skills of LEP students at end of project year* | 593 | 87 | 8 | 6 | 100 |
| English writing skills of LEP students at end of project year | 591 | 62 | 18 | 21 | 100 |
| Native language skills of LEP students at end of project year | 590 | 40 | 16 | 44 | 100 |
| Math achievement of LEP students at end of project year* | 594 | 64 | 16 | 20 | 100 |
| Science achievement of LEP students at end of project year* | 588 | 30 | 19 | 51 | 100 |
| Social studies achievement of LEP students at end of project year* | 589 | 29 | 20 | 51 | 100 |
| Grade retention or grade advancement* | 591 | 52 | 30 | 18 | 100 |
| School dropout* | 588 | 38 | 24 | 37 | 100 |
| Absenteeism/attendance* | 591 | 52 | 27 | 20 | 100 |

* Elements required by Title VII regulations

TABLE 43

(continued)

| Data Element | N | Described in 1989-90 Report | Collected But Not Reported | Not Reported | Total |
|---|---|---|---|---|---|
| Referral to or placement in special education* | 587 | 40% | 35% | 25% | 100% |
| Placement in gifted and talented programs* | 589 | 30 | 30 | 40 | 100 |
| Student enrollment in postsecondary institutions* | 588 | 9 | 11 | 80 | 100 |
| Academic achievement of former project participants* | 591 | 14 | 23 | 63 | 100 |
| Specific content areas of instruction being received by project students* | 590 | 56 | 20 | 24 | 100 |
| Clearly defined instructional methods being used in project (e.g., sheltered English)* | 592 | 73 | 11 | 16 | 100 |
| Specific instructional materials being used by the project* | 593 | 65 | 21 | 14 | 100 |
| Time spent on specific instructional tasks* | 589 | 48 | 22 | 30 | 100 |
| Academic preparation of project staff (degrees, certification, etc.)* | 588 | 66 | 25 | 9 | 100 |
| Teaching experience of project staff | 592 | 56 | 30 | 13 | 100 |
| Language capabilities of project staff* | 592 | 60 | 29 | 12 | 100 |
| Project-supported preservice or in-service training received by staff | 593 | 87 | 9 | 5 | 100 |
| Project-supported college or university training received by project staff | 594 | 71 | 14 | 15 | 100 |
| Activities to increase parent involvement | 593 | 83 | 10 | 7 | 100 |
| Activities to develop curriculum materials | 594 | 61 | 14 | 25 | 100 |

* Elements required by Title VII regulations

TABLE 44

Composite Measures of Evaluation and Report Completeness*
(N=594)

| Composite Score | Composite 1 | Composite 2 | Composite 3 | Composite 4 |
|---|---|---|---|---|
| 0-4 | 4% | 0% | 9% | 1% |
| 5-9 | 7 | 1 | 25 | 6 |
| 10-14 | 19 | 5 | 33 | 22 |
| 15-19 | 22 | 9 | 26 | 41 |
| 20-24 | 24 | 22 | 7 | 31 |
| 25-29 | 18 | 31 | (NA) | (NA) |
| 30-36 | 5 | 32 | (NA) | (NA) |
| Total | 100% | 100% | 100% | 100% |
| Mean | 18.7 | 26.0 | 11.8 | 16.8 |

* Composite 1 = Overall Completeness of Evaluation Report (possible range: 0-36)

Composite 2 = Overall Completeness of Data Collection (possible range: 0-36)

Composite 3 = Completeness of Report in Terms of Title VII Requirements (possible range: 0-24)

Composite 4 = Completeness of Data Collection in Terms of Title VII Requirements (possible range: 0-24)

We compared the composite completeness scores of different categories of projects. As Table 45 shows, projects initially funded in 1989, special alternative projects, and smaller projects all tended to have lower completeness scores. The differences between transitional projects and special alternative projects can be partially accounted for by their differences in size.

We also compared the composite completeness scores based on a broad range of evaluator characteristics. The results are presented in Table 46. In general, there were higher completeness scores when no one person had primary responsibility for the evaluation and when the evaluator had more training and experience.

TABLE 45

Mean Completeness Composite Scores* For Different Categories of Projects

| Category of Project | Composite 1 | Composite 2 | Composite 3 | Composite 4 |
|---|---|---|---|---|
| Initially funded in 1985 (N=127) | 19.0 | 26.1 | 12.2 | 17.0 |
| Initially funded in 1986 (N=82) | 19.5 | 26.6 | 12.6 | 17.4 |
| Initially funded in 1987 (N=66) | 19.5 | 27.1 | 12.6 | 17.7 |
| Initially funded in 1988 (N=128) | 18.9 | 26.6 | 12.0 | 17.1 |
| Initially funded in 1989 (N=191) | 17.6 | 24.9 | 10.8 | 15.8 |
| Transitional project (N=462) | 19.0 | 26.1 | 12.0 | 16.8 |
| Special altern. project (N=132) | 17.4 | 25.5 | 11.0 | 16.7 |
| Less than $100,000 (N=186) | 17.7 | 26.0 | 11.1 | 16.9 |
| $100,000- $149,999 (N=165) | 18.7 | 26.0 | 12.0 | 16.8 |
| $150,000- $199,999 (N=130) | 19.0 | 26.0 | 11.9 | 16.8 |
| $200,000 or more (N=113) | 19.8 | 25.9 | 12.6 | 16.6 |

* Composite 1 = Overall Completeness of Evaluation Report (possible range: 0-36)

Composite 2 = Overall Completeness of Data Collection (possible range: 0-36)

Composite 3 = Completeness of Report in Terms of Title VII Requirements (possible range: 0-24)

Composite 4 = Completeness of Data Collection in Terms of Title VII Requirements (possible range: 0-24)

TABLE 46

Mean Completeness Composite Scores* For Different Types of Evaluators

| Evaluator Type | Composite 1 | Composite 2 | Composite 3 | Composite 4 |
|---|---|---|---|---|
| Primary responsibility for evaluation- | | | | |
| External evaluator (N=403) | 18.7 | 25.6 | 11.8 | 16.5 |
| District evaluation staff (N=61) | 16.4 | 23.0 | 10.8 | 15.2 |
| Project director or other district staff (N=56) | 14.9 | 24.0 | 8.9 | 15.1 |
| No single person (N=84) | 20.1 | 27.6 | 12.8 | 17.9 |
| Highest academic degree- | | | | |
| Ph.D. (N=244) | 18.5 | 25.7 | 11.6 | 16.6 |
| Ed.D. (N=100) | 18.1 | 26.0 | 11.8 | 17.1 |
| Other (MA, BA, etc.) (N=182) | 19.0 | 25.2 | 12.1 | 16.3 |
| Major field of highest degree- | | | | |
| Education (N=267) | 18.9 | 26.0 | 12.1 | 16.9 |
| Psychology/sociology/ anthropology (N=139) | 17.9 | 24.3 | 11.3 | 15.6 |
| Evaluation/statistics/ measurement (N=65) | 17.4 | 26.5 | 10.9 | 16.9 |
| Other (N=53) | 20.1 | 26.4 | 12.8 | 17.1 |
| College courses in educational evaluation/ measurement/data analysis- | | | | |
| 0-7 (N=215) | 17.3 | 24.8 | 10.9 | 15.9 |
| 8 or more (N=270) | 19.2 | 26.1 | 12.2 | 16.9 |
| College courses about educating LEP students- | | | | |
| 0-2 (N=288) | 18.1 | 24.9 | 11.6 | 16.1 |
| 3 or more (N=200) | 18.7 | 26.4 | 11.8 | 17.0 |

* Composite 1 = Overall Completeness of Evaluation Report (possible range: 0-36)

Composite 2 = Overall Completeness of Data Collection (possible range: 0-36)

Composite 3 = Completeness of Report in Terms of Title VII Requirements (possible range: 0-24)

Composite 4 = Completeness of Data Collection in Terms of Title VII Requirements (possible range: 0-24)

TABLE 46

(continued)

| Evaluator Type | Composite 1 | Composite 2 | Composite 3 | Composite 4 |
|---|---|---|---|---|
| Additional training hours in educational evaluation/measurement/data analysis- | | | | |
| 0-70 (N=210) | 17.6 | 25.5 | 11.3 | 16.6 |
| 71 or more (N=219) | 19.4 | 25.9 | 12.3 | 16.7 |
| Additional training hours in educating LEP students- | | | | |
| 0-30 (N=223) | 17.7 | 25.2 | 11.4 | 16.3 |
| 31 or more (N=213) | 19.3 | 26.4 | 12.2 | 17.1 |
| Years involved in evaluating educational programs- | | | | |
| 0-16 (N=263) | 17.6 | 25.1 | 11.2 | 16.2 |
| 17 or more (N=263) | 19.6 | 26.1 | 12.4 | 16.9 |
| Years involved in evaluating Title VII projects- | | | | |
| 0-8 (N=255) | 18.0 | 25.3 | 11.5 | 16.4 |
| 9 or more (N=271) | 19.1 | 25.8 | 12.1 | 16.7 |
| Number of Title VII evaluations being done by evaluator- | | | | |
| 1-2 (N=220) | 16.3 | 24.9 | 10.3 | 16.2 |
| 3 or more (N=287) | 20.4 | 26.3 | 13.0 | 17.0 |

---

* Composite 1 = Overall Completeness of Evaluation Report (possible range: 0-36)

Composite 2 = Overall Completeness of Data Collection (possible range: 0-36)

Composite 3 = Completeness of Report in Terms of Title VII Requirements (possible range: 0-24)

Composite 4 = Completeness of Data Collection in Terms of Title VII Requirements (possible range: 0-24)

C. Dissemination of Evaluation Results

Yearly evaluation reports are a requirement of Title VII grants, and copies of those reports are supposed to be sent to the U.S. Department of Education. We were interested, however, in to whom else reports were distributed. Therefore, we asked with whom reports were shared at the local level and, to whom copies were sent at the federal/state levels. The results on these questions are shown in Table 47.

TABLE 47

Dissemination of Title VII Evaluation Reports
(N=608)

| Recipient of Report | Percentage of All Projects |
|---|---|
| Local | |
| Relevant school principals | 83% |
| District-level staff | 77 |
| School superintendent | 76 |
| School-level staff | 62 |
| Parents | 62 |
| School board members | 54 |
| Other members of the public | 19 |
| State/national | |
| OBEMLA project officer | 88 |
| U.S. Department of Education grants officer | 80 |
| State bilingual/ESL coordinator | 52 |

We also asked about discussions or meetings that were held in the district concerning evaluation results. Of those responding, 79 percent (N=594) said that they had held meetings or serious discussions with project staff and 61 percent (N=587) said that they had held meetings or serious discussions with school district officials. Table 48 shows the topics that were discussed in meetings with staff. Discussions of the meaning of the evaluation results and discussions of how the project might be improved were those most frequently cited.

We also asked who had attended meetings of school district officials in which Title VII evaluation results were discussed. As Table 49 shows, district staff, relevant school principals, and teachers were those most likely to be present.

TABLE 48

Topics Discussed in Staff Meetings About Evaluation Results
(N=608)

| Discussion Topic | Percentage of All Projects |
|---|---|
| What the evaluation results meant | 74% |
| How the project might be improved | 72 |
| Why the results turned out the way they did | 66 |
| What components of the project were most and least effective according to the results | 65 |
| Issues and problems in collecting evaluation data | 63 |
| How the evaluation plan might be revised | 42 |
| Other | 6 |

TABLE 49

School District Officials Attending Meetings About Evaluation Results
(N=608)

| Type of Staff | Percentage of All Projects |
|---|---|
| District-level staff | 46% |
| Relevant school principals | 43 |
| Teachers | 37 |
| School superintendent | 32 |
| Other school-level staff | 21 |
| School board members | 17 |

# VIII. USEFULNESS OF EVALUATION

## A. Perceived Usefulness of Evaluation

Evaluations of Title VII projects are meant to provide information for federal officials. They are also meant to provide information to local project and school officials for program improvement purposes. Therefore, we asked project directors and evaluators a series of questions about the usefulness of evaluations.

We first asked project directors to give an overall rating of the usefulness of the evaluation process for improving the project. The results are shown in Table 50. Most of the respondents describe the evaluation process as either "moderately useful" or "very useful."

TABLE 50

Usefulness of the Evaluation Process-
Ratings by Project Directors
(N=587)

| Rating | Percentage of Projects |
|---|---|
| Not at all useful - (1) | 3% |
| Of limited use - (2) | 11 |
| Moderately useful - (3) | 43 |
| Very useful - (4) | 43 |
| Total | 100% |

We compared the ratings of usefulness across different categories of projects and found no major differences based on year of initial funding, project type, and project size.

We also compared the ratings of usefulness based on a range of evaluator characteristics (education level, area of specialty, amount of training, experience with other Title VII evaluations, etc.), and found that only the type of person who had primary responsibility for the evaluation was related to ratings of usefulness. The highest ratings of usefulness were given when no single person was cited as having primary responsibility (mean = 3.4, N=82). Evaluations in which external evaluators had primary responsibility were given the next highest rating (mean = 3.3, N=391), while evaluations conducted by project or district staff were given lower usefulness ratings (mean = 3.0, N=113).

We asked both project directors and evaluators which elements of the evaluation process were most useful and least useful for program improvement. Eighty-five percent of project directors and 86 percent of evaluators listed at least one useful element, but only 33 percent of project directors and 42 percent of evaluators listed a least useful element.

There was a very wide range of response to both questions. We divided the responses into four major areas: (1) specific information concerning project outcomes; (2) specific information concerning project processes; (3) reporting and feedback; and (4) other elements. The elements most frequently mentioned as useful within these categories are shown in Table 51. The elements most frequently mentioned as "least useful" are presented in Table 52. The controversy surrounding standardized testing is illustrated by the fact that it appears on both lists.

B. Actions Taken as a Result of Evaluation

As another measure of the usefulness of Title VII evaluations, we asked project directors to list concrete actions, if any, which were taken in response to evaluation results in 1989-90. Two-thirds of the 608 respondents gave at least one response to this question, and the mean number of responses was 1.3. The most frequently provided responses are shown in Table 53. Improvements in data collection methods and improvements in staff training were most frequently mentioned.

We checked to see if the number of responses to this question was related to various evaluator characteristics (e.g., education, training, experience). The factor most strongly related to the number of responses was the type of person who had primary responsibility for the evaluation. Project directors reported the most actions taken when they (the project directors) had primary responsibility for the evaluation (mean = 1.5, N=47). When others had primary responsibility or if the responsibility was shared, the number of actions reported was lower (mean = 1.2, N=557).

TABLE 51

Elements of the Evaluation Process Rated as Most Useful

| Element of Evaluation | Project Directors (N=608) | Evaluators (N=560) |
|---|---|---|
| Information on project outcomes | | |
| Achievement test data | 9% | 30% |
| Data on criterion-referenced/local/teacher tests | 9 | 4 |
| Relationship of project implementation and student performance | 8 | 3 |
| English proficiency test data | 6 | 8 |
| Attendance/dropout data | 3 | 2 |
| Grade advancement/graduation data | 2 | 1 |
| Self-concept/school attitude data | 1 | 1 |
| Reclassification/exit data | 1 | 1 |
| Comparisons across years | 1 | 1 |
| Information on project processes | | |
| Perceptions of parents, teachers, staff, administrators, students | 13 | 21 |
| Progress towards objectives | 12 | 16 |
| Site visits/classroom observation | 11 | 19 |
| Assessment of staff development activities | 8 | 6 |
| Assessment of parent involvement activities | 5 | 3 |
| Assessment of program coordination/ management | 4 | 2 |
| Collection of student background information | 2 | 1 |
| Assessment of curriculum materials | 2 | 0 |
| Analysis of time in program/time on task | 2 | 1 |
| Evaluation reporting | | |
| Recommendations by evaluator(s) | 18 | 15 |
| Evaluation reports | 7 | 13 |
| Feedback to staff/teachers/project director | 6 | 20 |
| Documentation of outcomes for "outsiders" | 1 | 3 |
| Other | | |
| Clarification of project objectives | 10 | 10 |
| Information on project implementation from evaluator | 7 | 5 |
| Awareness of the need for documentation/ evaluation | 6 | 11 |
| Development of forms/procedures/data bases | 5 | 5 |
| Understanding evaluation/testing design | 5 | 8 |
| Identification of special student needs | 1 | 4 |

TABLE 52

Elements of the Evaluation Process Rated as Least Useful

| Element of Evaluation | Project Directors (N=608) | Evaluators (N=560) |
|---|---|---|
| Information on project outcomes | | |
| Standardized test data | 9% | 11% |
| Data on absences/tardiness | 1 | 2 |
| Data on retention/dropout | 1 | 5 |
| Comparison group results | 1 | 10 |
| Data on referral to special education | 1 | 2 |
| Data on students exited from program | 1 | 1 |
| Calculation of "relative growth indices" | 0 | 3 |
| Self-esteem/school attitude data | 0 | 2 |
| Information on project processes | | |
| Data on student backgrounds | 2 | 1 |
| Interviews/questionnaires with staff | 1 | 1 |
| Data on student time on task | 1 | 3 |
| Assessment of staff development activities | 1 | 2 |
| Data on teacher characteristics | 1 | 1 |
| Evaluation reporting | | |
| Graphs/statistics | 2 | 1 |
| Reporting requirements for first year projects | 1 | 1 |
| Lack of format or criteria from ED | 1 | 1 |
| Timeline for grant proposals (can't use report) | 0 | 2 |
| Other | | |
| Too much time collecting data | 4 | 4 |
| Insufficient resources for evaluation | 2 | 4 |
| Insufficient recordkeeping by project | 1 | 3 |

TABLE 53

Concrete Actions Taken in Response to Evaluation Results
(N=608)

| Action Taken | Percentage of Projects |
|---|---|
| Improved data collection/testing procedures | 24% |
| Improved staff training | 20 |
| Instructional activities revised/added/ reorganized | 18 |
| Revised curricular materials | 10 |
| Increased staff training | 9 |
| Revised parent involvement component | 7 |
| Revised project objectives | 5 |
| Classroom aides added/rescheduled | 4 |
| Revised evaluation design | 4 |
| Project activities institutionalized in school/district | 4 |
| Increased dissemination of project results | 3 |
| Replaced the evaluator | 2 |

C. Recommendations to Improve Evaluation

We asked both project directors and evaluators to suggest changes in the regulations and procedures of the U.S. Department of Education that would improve the evaluation process. Forty-one percent of project directors and 56 percent of evaluators gave at least one response to this question. The responses to the question were so diverse that we ended up coding 71 different responses. Those suggested changes fell into seven categories: (1) changes relating to technical design and instrumentation; (2) increases in the involvement of OBEMLA/ED; (3) selection/training of evaluators; (4) elimination of specific requirements; (5) increased emphasis on process/formative evaluation; (6) the timing of reports; and (7) other changes. The most frequent responses in these eight categories are shown in Table 54. The two most frequent recommendations across respondents were that more money should be devoted to evaluation and that less documentation should be required.

TABLE 54

Recommendations for Changes in the Regulations and Procedures of the U.S. Department of Education Relating to Evaluation

| Recommendation | Project Directors (N=608) | Evaluators (N=560) |
|---|---|---|
| Technical design and instrumentation | | |
| More innovative assessment (e.g.,portfolios) | 6% | 4% |
| Development of umbrella evaluation designs with common elements | 5 | 4 |
| Approve noncomparison group designs | 3 | 7 |
| Promote/distribute BEES | 1 | 3 |
| Drop support for gap reduction | 0 | 6 |
| Use electronic data collection | 0 | 5 |
| Increased OBEMLA/ED involvement | | |
| Project officers provide feedback on evaluations | 5 | 5 |
| A negotiated evaluation plan | 2 | 4 |
| Penalize projects that don't submit reports or don't follow their designs | 1 | 6 |
| Selection/training of evaluators | | |
| Train a cadre of "good evaluators" | 3 | 1 |
| Evaluator should not be district employee | 2 | 1 |
| Lack of format or criteria from ED | 1 | 1 |
| Elimination of requirements | | |
| Require less documentation | 5 | 8 |
| Delete postsecondary enrollment data requirement | 1 | 1 |
| Delete special education/gifted data requirement | 0 | 2 |
| Increase emphasis on formative evaluation | | |
| Less emphasis on standardized testing | 3 | 3 |
| Support formative evaluation | 1 | 4 |
| Require project directors/principals to do process evaluation | 1 | 4 |
| Timing of reports | | |
| Stagger reporting periods for different types of data/ require longer term studies | 1 | 3 |
| Clarify when reports are due | 0 | 6 |
| Other | | |
| More money for evaluation | 4 | 9 |
| Increase points on application for evaluation | 1 | 1 |
| Have lower expectations for small/rural projects | 0 | 3 |
| Require an evaluator assist in application | 0 | 2 |

The last question that we asked of both project directors and evaluators was whether there was any evaluation-related assistance that they would like to receive from the U.S. Department of Education. The most frequent responses to this question are presented in Table 55. Additional workshops on evaluation and a clear summary of requirements were mentioned most frequently overall.

TABLE 55

Assistance Related to Evaluation Requested of the U.S. Department of Education

| Type of Assistance | Project Directors (N=608) | Evaluators (N=560) |
|---|---|---|
| Feedback on evaluation reports | 7% | 2% |
| Model evaluation reports | 6 | 4 |
| National/state/local workshops on evaluation | 5 | 8 |
| Clear summary of guidelines/requirements | 4 | 9 |
| Simple evaluation designs | 3 | 1 |
| More money for evaluation | 3 | 5 |
| Training in alternative assessment techniques | 3 | 3 |
| Information from results of other projects | 3 | 2 |
| EAC contact | 2 | 3 |
| Guidance to project designers on evaluation | 2 | 1 |
| A standardized data base | 2 | 1 |
| Updated version of BEES | 1 | 3 |
| Send evaluators to OBEMLA Management Workshops | 0 | 6 |

In summary, project directors and evaluators had a wide range of suggestions for improving the Title VII evaluation system, and many of them requested additional assistance so that they could improve the evaluation process.

# APPENDIX 2

"Summary of the Case Study Findings"

(Chapter III from The Case Study Report)

## III. SUMMARY OF THE CASE STUDY FINDINGS

In this next section of this report, we review the findings of the 18 case studies to provide a summary overview of the evaluation activities and their strengths and weaknesses. Title VII projects and their evaluations are directed first of all toward examining the effect of the projects' effort on student outcomes. Therefore, in Section A the review will focus on the quality of the summative evaluations that were carried out in the case study sites. Next (Section B), the additional process evaluation activities that were found to be carried out are briefly described. Finally, since the central objective of Title VII projects is to increase the capacity of school districts with respect to the instruction of limited English proficient students, Section C will address the question of capacity-building and the role played by the evaluations in promoting institutionalization of the improvements introduced through the projects.

### A. Summative Evaluations

According to the Title VII regulations, all Title VII-funded local projects are required to evaluate the success of the projects in increasing student academic performance. Projects are required to provide data on the project activities and on student variables such as time in the project, grade retention, and absenteeism, as well as student performance data based on standardized tests of academic achievement. In this section we examine the extent to which the case study projects collect and analyze data as called for in the regulations. In addition, we review the projects' use of the Bilingual Education Evaluation System (BEES), which was developed to serve as a guide for Title VII projects in carrying out their evaluations.

#### 1. Collection/reporting of required student/project data

Table 2 presents a summary of the extent to which 24 specific data elements required by the Title VII regulations were included in the evaluation reports for 1989-90 submitted by the 18 case study projects. This table provides an overview of the relative completeness of the evaluations based on the project directors' responses to the survey questionnaire.

Overall, in terms of the required elements for the Title VII evaluations, the case study projects varied considerably in their degree of consistency with the Federal requirements, i.e., in the extent to which the data that are required in the regulations were collected and reported. As Table 2 demonstrates, none of the case study sites included all of the required elements and, in fact, one site's evaluation report for 1989-90 did not include any of the required elements. However, a few of the projects did provide much of what is outlined in the regulations. For example, the evaluations carried out in sites #13 and #14 generally fulfilled all of the evaluation requirements (although there were limitations in the achievement test data that were presented in site #14).

TABLE 2

Extent to which Specific Data Elements Required by Title VII Regulations Are in Title VII Project Evaluations for 1989-90

| | Case Study Sites | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Data Element | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| Years of previous education of project students | | • | | | | | | • | | | | • | | | | | • | • |
| Oral English skills of LEP students at beginning of project | | • | | • | • | • | | • | • | • | • | • | • | • | • | • | • | • |
| English reading and language skills of LEP students at beginning of project | | • | | • | • | • | • | • | | • | • | | • | • | • | • | • | • |
| Academic achievement in math of LEP students at beginning of project | | • | | • | • | • | • | • | | • | • | | • | • | • | • | • | • |
| Academic achievement in science of LEP students at beginning of project | | • | | • | • | | | • | | | • | | • | • | | | | |
| Academic achievement in social studies of LEP students at beginning of project | | • | | • | | | | | | | • | | • | • | | | | |
| Oral English skills of LEP students at end of project year | • | • | | | • | • | | • | | | • | • | • | • | | • | • | • |
| English reading and language skills of LEP students at end of project year | • | • | | • | • | • | • | • | • | | • | • | • | • | • | • | • | • |
| Math achievement of LEP students at end of project year | • | • | | • | • | | • | • | | | • | • | • | • | • | | • | • |
| Science achievement of LEP students at end of project year | | • | | • | • | • | • | • | | | • | | • | • | | | | |
| Social studies achievement of LEP students at end of project year | | • | | • | | | • | | | | • | | • | • | | | | |
| Grade retention or grade advancement | | • | | | • | | | | • | | | | • | • | | • | • | |
| School dropout | | • | | | • | | | | | | | | • | • | | | • | |
| Absenteeism/attendance | | • | | | • | | | | | | | | • | • | | • | • | |

TABLE 2
(continued)

| | Case Study Sites | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Data Element | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| Referral to or placement in special education | | • | | | • | | | | | | | | • | • | | • | • | |
| Placement in gifted and talented programs | | | | | • | | | | | | | | • | • | | • | | |
| Student enrollment in postsecondary institutions | | | | | | | | | | | | | | | | | | |
| Academic achievement of former project participants | | | | | | | • | • | • | | | | | | | | | |
| Specific content areas of instruction being received by project students | | | | | | • | • | • | • | | • | | • | • | | | • | • |
| Clearly defined instructional methods being used in project (e.g., sheltered English) | | | | | • | • | • | • | • | | • | | • | • | | | • | • |
| Specific instructional materials being used by the project | | | | | | • | • | • | • | • | • | | • | • | | | • | • |
| Time spent on specific instructional tasks | | | | • | • | • | | | • | • | • | | • | • | | | | |
| Academic preparation of project staff (degrees, certification, etc.) | | • | | • | | | • | • | • | • | • | | • | • | | | • | • |
| Language capabilities of project staff | | • | | | | | • | • | • | | • | | • | • | | | | • |
| Percent of Required 24 Elements | 12 | 71 | 0 | 46 | 62 | 42 | 50 | 62 | 42 | 25 | 67 | 21 | 88 | 88 | 21 | 38 | 62 | 50 |

In some of the sites where about half of the required data elements were provided, the evaluations carried out were strong in other regards. For example, in site #7 the evaluation made use of the gap reduction model (as outlined in the BEES) for examination of student achievement outcomes and provided follow-up data on former LEP students; the evaluation report also included discussion of reliability of the data obtained. In addition, the project carried out formal process evaluation activities which were reported in the evaluation report. However, data were missing for several required elements. Data on oral English proficiency outcomes, or on school attendance were not collected or reported and other areas such as staff development, parent involvement were not described.

Site #18 also provided about half of the elements described in the regulations. Several analyses of these data were carried out supported by a strong computerized data base within the district. Analyses of student test score data (achievement tests, English proficiency tests) were carried out providing comparisons across time not only by grade and school but also by language group and English proficiency level. In addition, changes in components of the instructional services (such as time in the language lab, after hours classes, etc.) were documented. The project was also careful to separate out students according to the length of their participation in the project.

The budgets for evaluation in the projects were generally low, most often less than $5,000, with several having only $3,000-3,500 available for evaluation. Especially when projects are strong in some areas of their evaluation, it may be that limited resources prevent their being able to provide all of the other data and analyses that are required.

Also, the collection of some variables outlined in the regulations is difficult for projects and therefore the variables are not included in the evaluations. For example, in site #1, the report did not include many of the variables simply because the data were not available within the district. The report provided to the Department of Education did not include, for example, data on attendance, dropout, retention, or referrals. In cases such as this, part of the difficulty in obtaining data is due to district methods for data collection that do not permit isolation of the data for students included in the project. As an example, in site #14, SES data were available but data for the project students could not be disaggregated from among these. In site #6, the evaluator had difficulty in obtaining the student data needed. With no districtwide student data-base system, data on attendance, retention, classroom grades, and so forth were kept in a number of different locations. In addition, as the project director reported, some of the district staff had a "proprietary" feeling about some of the data. In this project, the director eventually created a separate project data base as a solution to the reporting problems.

In some cases, the required data are available but simply are not recorded or maintained for the project records and, therefore, are not included in the evaluation report. In site #11, the evaluation report indicated simply that objectives in terms of

absenteeism and tardiness were achieved without presenting any data. Also, while teachers provided a midyear report on the most typical schedule showing time on different types of tasks, no discussion of student time on specific tasks was included in the report. In site #8, as well, data on several of the student and program variables were available and on file but were not reported in the evaluation.

The situation in site #3, where none of the 24 required elements were present, was complex. This was a site in which the gap reduction approach had been utilized in the earlier years of the proejct. However, turnover in staff, shifts in project goals, apparently some lack of support from the district testing office, plus some errors in data entry resulted in a poor evaluation report for 1989-90.

In summary, the student data variables that are included in the regulations appear to be difficult to provide for many, if not most, of the projects. Data on students, such as attendance, time on specific instructional tasks, amount of time in the program, oral English language proficiency, dropout, grade-retention, and absenteeism were not consistently collected and reported. When these data were included, the reports often did not provide pre-post comparison data or data comparing the project students with other groups. In some cases, the data on variables, for example, attendance, were available within the district or school but were not included within the report. In part, the lack of reporting on these data is due to the data collection burden required in separating out project students from other students in the school and from the lack of a comprehensive student database system in the local district.

2. Use of the BEES (Bilingual Education Evaluation System)

The Bilingual Education Evaluation System (BEES) was developed through Department of Education funding to provide assistance to Title VII projects in carrying out their evaluations. The BEES outlines suggested means of fulfilling the requirements for evaluation. With regard to the measurement of student achievement outcomes, it includes description of a "gap reduction" model in which the outcomes are measured through comparison of changes over time in the relationship between test scores of project students and national norms or the scores of a comparison group.

Although the BEES has been available for almost five years, not all of the case study projects were familiar with it and levels of familiarity varied considerably. In some locations neither the project director nor the evaluator were familiar with the BEES. For example, this was true for site #8; also, in site #9, neither of the evaluators was familiar with the BEES and the teacher coordinator who effectively served as project manager had no background in evaluation (and therefore presumably no familiarity with the BEES). In another site, the project director was not familiar with the BEES while the evaluator was familiar with it (indicating that for the purposes of the evaluation there was no discussion of the BEES and its application to the project

between the evaluator and director). However, in most of the case study sites, the evaluator and project director were familiar with, or at least aware of, the BEES.

Familiarity with the BEES did not always guarantee its use and project staff were not always in agreement with the emphasis on student achievement outcomes that the BEES represents. The comments of the evaluator and project director in site #15 were similar to those of staff in other projects. Although they were familiar with the BEES and found it moderately useful, they believed that the use of the standardized tests were not the most appropriate measure of project outcomes. In another site (site #14), where both the evaluator and the project director were familiar with the BEES, they stated that, although they believed it overall to be of limited use, the gap-reduction portion of the BEES was the most useful. (This was a project that used the gap-reduction approach.)

Although the gap reduction model was specifically referred to in slightly fewer than half of the 18 case studies, these did not always reflect actual implementation of the model. For example, in site #3, there had been a dedicated effort to utilize the gap reduction model in earlier years of the grant but, with personnel changes and shifts in the focus of the project, the gap reduction model was not continued through all years of the project. Similarly, in site #2, a gap reduction approach had been utilized in earlier years, but the current evaluation director for the project believes that the approach outlined in the BEES is unneccessarily elaborate and not cost effective. In site #13, the evaluators and the project coordinator were familiar with the gap reduction model but made the determination not to use it due to the level of effort required as compared with the expected benefit to the project. The gap reduction model was in use in site #5 but was used in conjunction with an ongoing process evaluation, which was felt to be much more useful to the project compared to the achievement test outcomes. In site #4, the evaluation report stated that a gap-reduction model was being used but the data presented were not completely consistent with the model. In this site, the "gap-reduction" analysis focused on data for a single point in time rather than examining differences between student test scores and the national norm over more than one point in time. The gap-reduction approach was apparently used appropriately in sites #14 and #16.

Local conditions make some aspects of the Federal requirements very difficult, if not impossible, to implement in certain cases. For example, in site #1 (a small, rural district), the high turnover in students precluded carrying out any analyses utilizing comparison groups although this had been proposed as part of the original evaluation plan (and, given the nature of the student population should have been recognized as a difficulty in the planning stages of the project). Similarly, where there is much mobility in the school or district, it is difficult and complex to maintain records on length of participation in order to select those students who have the requisite number of days to be considered as participants in the project (although site #18 was able to do so through the use of the district's comprehensive data base).

In summary, there was familiarity with the BEES for most of the case study projects. Although several of the projects were familiar with the gap-reduction model, not all of these used it in the evaluations. A frequent comment from many of the projects, including those who used the gap-reduction model, was that they were dissatisfied with the reliance on the use of standardized tests as measures of project outcomes.

3. Validity and reliability of data

Where achievement data or other data are presented in project evaluations, these data are not consistently linked with participation in the project improvements. Therefore, the outcomes reported are not valid measures of the project's effect. For example, in site #3, the outcome data presented were not linked to students who had actually been exposed to the project's instructional approach; in fact, it was not known whether or not these students had received instruction within a classroom in which the approach was being utilized. This lack of relationship between test scores analyzed and actual implementation of the intervention was also a problem noted in site #6. It is likely that this is a problem in other projects as well, particularly in projects where documentation of project classroom implementation is not carried out.

Reliability of data is an additional issue. For example, in some cases, as in site #7, reliability is a concern and the reliability of the data was given a clear examination, in other cases, there is a need for further assistance in developing reliable data collection methods. In site #9, the collection of student data was carried out by a teacher without a background in evaluation who carried out the data collection without any assistance from the project evaluator or district staff. As a result, there was no standardization in the collection of the data, the process was not formalized, and the data were collected haphazardly, weakening the evaluation. In other cases (e.g., site #11 among others), tests were administered to students at different times or, as in one case, it was not clear which form or level of the achievement test was actually administered to students.

In summary, problems with validity and reliability were present in the evaluations and called into question the usefulness of the data as measures of project outcomes. The problems observed indicate a need for further attention to these issues.

B. Process Evaluations

The Federal requirements do not include reference to process or formative evaluation. However, despite this fact, almost all of the case study sites carried out process evaluation activities. There was for some case study projects a substantial effort devoted to the conduct of process evaluation activities, although in most cases, the process evaluation findings were not included in the evaluation report submitted to the Department of Education.

The efforts devoted to process evaluation apparently have been motivated by the need for timely data that would assist the project staff in making decisions over the course of the implementation, and by dissatisfaction with the evaluations as required by the regulations. Staff at several projects were unhappy with the reliance on student achievement test outcomes as measures of project success; they believed these were not appropriate for their students and that other outcome measures would be more valid indicators of the effect of the project's efforts. In addition, staff at several sites commented on the late receipt of the formal evaluation report as precluding use of the findings to the next year's planning (e.g., in site #13 the first year's evaluation report was not received until close to the end of year 2). Therefore, projects apparently turned to process evaluations to fulfill their need for more valid and timely indicators of how well their projects were progressing. The findings developed through ongoing process evaluations were felt to be more useful since, based on the information gathered in the process evaluation activities, project staff could make better, more informed decisions regarding adjustments and modifications in the implementation of the project.

Some quite well-developed process evaluations were carried out in sites where the completeness of the evaluations in terms of required data elements was not very high. For example, project staff in site #5 devoted considerable effort to process evaluation activities which were in addition to a well-defined and implemented summative evaluation that included use of the gap reduction model. In this site, the staff designed and utilized a formative evaluation system to monitor the development of the science curriculum being implemented and to monitor other aspects of the project. No information on this extensive effort was reported in the evaluation report; only the required summative data were reported. Also, a very well-defined process evaluation was carried out in site #16, where a specific instrument for monitoring implementation of the project was used, and in site #18, where a strong process evaluation involved reviews of the project that were carried out on a monthly basis. As shown in Table 2 above, the evaluation reports for these projects included only 62 percent, 38 percent, and 50 percent of the required data elements. Thus, it may be that in working with very limited resources (evaluation budgets of $3,500 for site #5, $4,500 for site #18, unknown for site #16) and a clear sense of their evaluation needs, the projects determining that a portion of their resources should be applied toward the process evaluations, even if that means less than complete compliance with the regulations.

In sites #5, #16, and #18, the process evaluation activities incorporated formalized procedures and documentation. However, not all of the process evaluations were as structured or well-documented. Frequently, the process evaluation activities were more informal and not well-documented and, therefore, it is hard to judge overall the range and quality of the process evaluations that were carried out.

In summary, the projects instituted process evaluation activities to obtain immediate feedback on how well the project was progressing and to obtain information that could guide decision making in the course of implementing the project. Due to the general lack of documentation of process evaluation for most of the projects, it is difficult to assess how well these activities are addressing specific components of the project. However, project staff reported that the process evaluation activities were more useful for them than the

summative evaluation activities because they could be utilized in project decision making and better reflect the progress of the projects. In addition, in those projects where less process evaluation was occurring than in other sites, project staff indicated that they were interested in incorporating more process evaluation into their projects (e.g., site #11, site #13). Also, where process evaluation activities were minimal, limited evaluation resources apparently led to the decision to focus on the required evaluation activities in order to comply with the regulations.

Fullan (1985) points out the importance for any information gathering to have specific procedures in place for the use of that information. This was true for the case study sites, where the process evaluation findings were valued for their immediate application to decision making. Similarly, Purkey & Smith (1985) have noted that the late arrival of evaluation data at a school precludes meaningful use of the data and results in the data losing meaning and value for those working on the improvement. Project directors' complaints about the late arrival of evaluation reports emphasized the fact that they arrived too late to be of use in planning for the next year. In the next section, evaluation activities are discussed further, in relation to their potential for assisting in the projects' capacity-building.

## C. Capacity-Building and Title VII Evaluations

The regulations for the Title VII project evaluations do not address capacity-building. However, since the purpose of the Title VII program is to promote capacity-building, this is an aspect of project implementation that is of central importance. For this reason, we reviewed the literature on change and improvement processes within schools/districts to gain perspectives on how change is facilitated and successfully implemented and to understand the role of evaluation in this process. We then examined components of the case study evaluations in relation to the change research findings to assess how well, if at all, the evaluation activities that are carried out support the institutionalization of the improvement within the local district and schools.

To begin, it should be noted that evaluation can be related to capacity-building in two ways. First, evaluation activities can serve as part of the capacity-building process, through examination of the progress in implementation of the project. Second, there can be evaluation of capacity-building as a specific outcome, to examine the degree to which the project has become institutionalized or has taken the necessary steps toward development of institutionalization of the improvement. In this section, we discuss both of these aspects of evaluation related to the development of capacity-building within Title VII projects.

### 1. Evaluation as a part of the capacity-building process

First, the pervading dissatisfaction (referred to in sections A2 and B above) among Title VII projects with the current reliance in the regulations on formal summative evaluation is consistent with the literature's emphasis on evaluation activities that answer local project needs for information about the status of implementation. That

is, local projects recognize that they need information that tells them how well the implementation is proceeding and where adjustments are necessary. This type of information is not provided through the evaluation requirements as currently specified in the Title VII regulations since these focus on addressing the Federal government's need for information on project/program effectiveness.

The literature on change and improvement processes suggests that the development of new capacities within the school or district relies on information-gathering which addresses the local project's need for data regarding the implementation process and activities. That is, it is essential that information-gathering efforts result in findings that can be utilized by the project staff to improve the effectiveness of the implementation effort. The implication of this is that in addition to the monitoring of student outcomes, other aspects of the improvement effort require attention. These include:

Elements related to implementation:

- the nature of decision making (e.g., what is the process, who are participants in decision making, what are the roles and relationships among participants);
- the nature of the staff development process (e.g., nature and amount of follow-through, degree of support for teachers as learners);
- the extent/quality of implementation of the improvement effort (e.g., to what extent are staff utilizing the new approach in the classroom, how well is the new approach/materials being implemented);

Elements related to the implementation context:

- the existence and level of support for the improvement effort within the local context (schools/district);
- the existence and degree of "fit" of the improvement effort within the local context (school/district);

Elements related to the evaluation of the improvement effort:

- information gathering/evaluation that is useful to those involved in the change effort; and,
- dissemination of the findings from the evaluation.

Below, the findings of the literature on change and improvement processes are presented with regard to each of these aspects of a change effort. The findings are then related to the components of the Title VII case study projects. The purpose of this comparison is to examine the extent to which the Title VII project activities support the development of change, that is, the extent to which projects' activities are directed toward effective capacity-building.

a. Elements related to the process of implementation.

While much effort has been given to discussion of what makes instruction for limited English proficient students effective, less effort has been given to how to put improved approaches into place. This weakness parallels a weak point noted within the general school effectiveness research, which listed characteristics of effective schools but did not give any insight into how the effective schools came into being. That is, it said very little about the process of change (Fullan, 1985; Sparks et al., 1985). More recently, efforts in school reform have begun to focus on identifying how improvement can be most effectively introduced and sustained at the local level. For the most part, the local school has now become the primary target for change (Purkey & Smith, 1985; Wood, Freeland, & Szabo, 1985), out of recognition that the local factors determine the nature, amount and pace of change (McLaughlin, 1990).

One central theme found within the literature on change and improvement processes is the need for "restructuring" of decision making and staff interactions so that all participants take on "ownership" of the change effort. This collaborative effort must be supported by a school "culture" in which there are shared norms and goals and honest and open communication (e.g., David, Purkey, & White, 1989; Deal, 1990; Goldman & O'Shea, 1990; Lieberman & Miller, 1990; Simpson, 1990; Miles & Louis, 1990, Saphier & King, 1985; Glickman, 1991). In addition, there is an emphasis on the nature of the staff development process as the key to effective implementation of change. This emphasis goes in hand with monitoring the degree to which the improvement effort is actually being implemented. Both the nature of decision making and the nature of staff development activities as described in the change process literature have implications for Title VII projects. These are discussed below.

(1). **"Restructuring" of decision making.**

The restructuring of decision making places more of the decisions at the local school level, where there is an understanding of the local conditions that will affect implementation. Restructuring also involves all participants in collaborative decision making. That is, implementation is more likely to be successful when all who are involved in the change effort feel a sense of ownership of the change

(Tillema & Koster, 1990). The implication of this finding is that teachers become partners in the change effort to a much greater degree than in the past; they share in decision making about the nature of the change and its implementation. This changes the principal's role--or, in the case of Title VII projects--the project director's role from that of a strong and directive leader to a leader of leaders or a "coordinator of teachers as instructional leaders" (Glickman, 1991).

Within the case study sites reviewed in this study there was one site in which teachers were very active participants in decision making. In site #18, teachers and administrators were involved in all aspects of the project's decision making. In other cases, e.g., site #6, teachers and school staff were aware of the project's evaluation plan although it is not clear what level of input they had into the project. In general, it is not clear what level of decision-making involvement teachers had in the case study sites. However, with regard to evaluation specifically, teachers' roles were frequently limited to providing data on their students and their knowledge of the overall evaluation plan was similarly limited.

(2). **The nature of the staff development process**

The effectiveness of a project hinges on the effectiveness of staff development efforts, since teacher change is central to any improvement effort. To be effective, staff development must be an ongoing process that continues to involve the participating teachers and to assist them through all stages of the project's implementation. An ongoing process responds to the different needs that exist at different points in the implementation of a change effort. Thus, according to the research on change and improvement, teacher change and the continued use of the improvement being introduced is more likely to occur when (a) there is as much assistance to staff as possible during the early stages of implementation using a variety of formats (e.g., meetings, peer coaching, workshops) (Fullan, 1985); and, (b) there is follow-through over the course of the implementation that addresses teachers' needs at each stage of implementation and that involves teachers through active involvement in the design and implementation of project activities (McDonnell, 1985).

Follow-through in staff development is most effective when it incorporates a recognition that teacher change is an incremental, developmental process that requires time and opportunities for practice and feedback. (Fullan, 1985; Huberman & Crandall, 1983). That is, it is necessary to provide opportunities for trials in classrooms followed by further discussion and specific on-site assistance in developing use

of the model (Fullan, 1985; Huberman & Crandall, 1983; Guskey, 1985; Tillema & Koster, 1990).

Also, (c), part of staff development is the process of assisting teachers in defining new roles as participants in decision making and ongoing improvement efforts. Since teacher change occurs more when teachers are actively involved in and aware of the need for the change, then it is important for teachers to be familiar with the research related to the improvements being made. Thus, there is an emphasis being placed upon "teachers as learners" (Fullan et al., 1990) and on the development of schools as "universities for teachers," reflecting a focus on the need for teachers to understand research and other knowledge related to the school's (or project's) improvement goals and efforts (Richardson, 1990). This is promoted by giving teachers information and by giving them time and opportunity to seek out and explore relevant information. In this way, teachers develop not only skills but also an understanding of the improvement effort that can guide them in their work and help them contribute to the further definition of the improvement.

In summary, teacher change is promoted through an ongoing process that incorporates assistance and encouragement to teachers as they learn to implement the new method/approach, and that promotes the further development of teachers as informed professionals. These types of staff development activities lead to the development of both the "will" and the "skill" (Miles & Louis, 1990) necessary for implementing the change.

Although almost all of the case study projects include some mention of staff development in their evaluation reports, these generally provide very little in-depth information on the nature of staff development efforts or processes. For example, in several cases, the discussion of staff development in the evaluation reports was limited to the listing of workshops provided to the teaching staff (e.g., sites #1, #5) or to mention of the credits earned through university courses (e.g, site #11). While teachers in some sites (e.g., sites #3, #10, #15) evaluated the workshops they received and even suggested areas of need for future workshops, follow-through with teachers subsequent to the workshops was not apparent in most cases. In site #18, however, teachers completed monthly self-assessments related to the project's implementation; this self-assessment was in combination with apparently regular staff interaction/discussion, which represents some follow-through on implementation.

Based on the limited information available on staff development efforts within the case study data, it appears that staff development efforts are

not carried out as ongoing, developmental processes. Rather, staff development more frequently appears to consist of presentation of workshops/inservices without much emphasis on follow-up assistance to teachers in putting the new practices into use in their classrooms. However, this apparent lack of focus on follow-through for the staff development efforts is a serious weakness, since the literature shows that teacher change is central to the success of an improvement effort. The research on change processes suggests, therefore, that staff development is an area that requires further attention within the Title VII projects and that the evaluation process can make a substantial contribution in this regard.

(3). **Extent of implementation.**

Closely related to the issue of follow-through with teachers is the question of the extent to which an improvement is actually being put into practice, and the degree of success with which it is being implemented. In assessing the outcomes of an improvement, it is critical to have this type of information regarding the actual use of the improvement. As noted earlier, not all projects have a clear enough idea of how broadly and how well the improvement they are promoting is actually being put into place. That is, they are not able to identify which students are actually affected by the improvement in order to clearly examine the project's effect on student outcomes. In sites #3 and #6 there was no clear linkage of student outcomes with actual use of the improvement. Actually, in site #3 there was some attempt to measure implementation in a very rough way, i.e., through use of the project materials (check-outs from the project library). However, this indirect measure of project implementation was not used to link implementation with student data. In site #9, there was no documentation of what was taught or of methods used with students.

Although implementation was generally not assessed, at least one site (site #16) did focus directly on implementation, through the use of a "Project Implementation Evaluation." Also, site #18 carried out a regular "Quality Review" that may have also served the purpose of examining degree of implementation. In other projects, classroom observations were carried out (e.g., #4, #5, #6, #7, #8, #15). The findings of these were not reported in the evaluation reports and, therefore, the focus of these observations (including whether they were used to assess degree of implementation), is not clear.

b. <u>Elements related to the implementation context</u>.

As mentioned earlier, the current focus in reform and change efforts is on the individual school, with an emphasis on decision making carried out by school

staff. At the same time, however, there is a clear awareness of the fact that the school cannot be considered as separate from the local district context and, in fact, depends on the district for support. The development of a collaborative and supportive relationship that involves "active engagement" between the school staff and district personnel is very important for the success of a change effort within a school (Fuhrman, Clune, & Elmore, 1988; Purkey & Smith, 1985; Miles & Louis, 1990). Although schools need room to make local decisions, they also need assistance with problems and support in carrying out an improvement program (Huberman & Miles, 1984).

For local Title VII projects, the implementation context includes relationships with individual schools as well as with the district. Often, the program for limited English proficient students is seen as apart or separate from the regular program. However, the change literature describes improvement efforts as being more successful when they represent a systemic change as opposed to a piecemeal approach, for instance, as involving the whole school rather than narrow parts of the school program (Purkey & Smith, 1985). For Title VII projects, the implication of this recommendation is that the projects should attempt to at least inform other school staff and to involve them in the improvement effort as much as possible.

Thus, the literature on change and improvement processes suggests that for the improvement effort to be sustained (1) there must be support from the local context, i.e., the school and the district. In addition, (2) this support is more likely and success of the change effort is more likely when the improvement is perceived as "fitting" within the local context, for instance, as meeting existing needs and as being consistent with existing policies and practices.

(1). **The level of support within the schools/district.**

The level of support present in the school/district is an aspect of change process that is perhaps particularly important for local Title VII projects. In many cases they exist apart from the regular program, they may be conducted in schools in which the principals are not familiar with the activities or goals of the project, and the teachers involved may operate in isolation. However, research on change processes shows that teachers change their behaviors more in schools where the principal is supportive (Stallings & Mohlman, 1981; McLaughlin, 1990; Berman & McLaughlin, 1978) and where there is a "culture" that supports the change.

In some of the case study sites, the Title VII project did not have much support from principals. For example, in site #3, the project manager mentioned problems in communicating with teachers since principals did not always ensure that letters to the project teachers were

distributed as requested. In this site, principals frequently did not have any knowledge of the project's activities or knew simply that a couple of the teachers in the school were doing something new. In other sites, principals or assistant principals contributed to the design of the evaluation of the project (e.g, sites #1, #2, #7, #8, #13) while in still others, the principals were aware of the project and had some familiarity with the evaluation (e.g., site #5, #14), although to varying degrees. For some cases, principals had very little or no awareness of the project's evaluation (e.g., sites #10, #12), or familiarity varied by school (e.g., site #11). Although familiarity with the evaluation plan cannot be directly equated with familiarity and awareness of a project and support for it, the various levels of involvement/familiarity with the projects on the part of the principals suggests that Title VII projects are not consistently operating within supportive local school contexts.

Similarly, the project may not have the full support of the district and may lack important assistance because of this lack of support. For example, in district #3, the district context began to shift toward less support of the approach taken in the project, leading to consequent shifts in the focus of the project. In districts where the central district testing office is not involved at all with the development of the evaluation plan for the Title VII project, then there is some indication that the project is considered outside of its scope and that the district is not fully supportive of the project. For example, in site #4, the district's evaluation and testing staff did not have any input into the intial evauation design and, other than providing aggregated test scores as needed, did not provide any assistance in the implementation of the evaluation. This was also the case in site #3. Thus, the patterns found in the case studies indicate that for several of the projects support for the change at the school and district level may be lacking. If so, then the research on change processes suggests that this is an area that deserves further concern.

(2). **The "fit" of the project within the school/district context.**

Related to the concern for the context in which the Title VII projects operate is the concern for the fit of the project within the school and district. Research has demonstrated that changes are more likely to be put into practice and to remain as part of instructional practice when those involved acknowledge the need for the innovation (Roberts-Gray & Gray, 1983) and the change being implemented has been developed to address that local need (Berman & McLaughlin, 1978). Also, change is more likely when the new practices are congruent with the teachers' present philosophy and practices (Guskey & Sparks, 1991); presumably, this is also true in relation to existing philosophies within schools and districts.

Therefore, if an innovation is not perceived as meeting a specific district- or school-level need, then it is not likely to be implemented to any consistent degree (McDonnell, 1985). If teachers do not perceive the innovation as meeting identified needs, and if they and other staff and administrators see the project activities as not fitting well within the overall program, then the project will be weakened and chances for continuation bleak. These findings imply that in implementation of a change effort, it is critical to look at the impact of the project on other aspects of the school or district program and its relationship to existing philosophies/practices to identify problems that should be addressed. If the "fit" of the improvement within the school and district context is not maximized, then chances for its success will be decreased.

Within the case studies, problems of "fit" within the local context were observed and in some cases caused serious difficulties in implementation. For example, in site #9, a shift in district policy to a program of parental choice made the implementation of the project as originally planned almost impossible. The project had to work with several different schools over the course of its implementation, weakening its original plan for consolidating students to work more effectively with them. Similarly, in site #1, a shift in district policy away from the use of native language support in instruction and toward more of an emphasis on English instruction, undermined the project's efforts in use of native language for instruction. These examples indicate that the concept of "fit" is one that can have important implications for the long term success of a project's improvement effort.

c. Elements related to the evaluation of the improvement effort.

Examination of the progress and outcomes of an improvement effort is a critical component. Information must be gathered to assess the activities being carried out and the outcomes of the improvement effort. However, as Fullan (1985) emphasizes, evaluation information should be linked explicitly to procedures for utilizing the information so that the information can be used in furthering the improvement effort. The necessary first step toward utilizing the information is its dissemination.

Second, in order for a project's efforts to be sustained, program improvement efforts should be viewed as a catalyst for continued actions directed toward addressing the needs of the students and school. There must be a commitment to ensuring that there is an ongoing process that allows the school or schools to continue the improvement effort. In the absence of deliberate measures to build continuation of or institutionalization of the improvement, the natural forces of attrition will lead to its disappearance

(Fullan, 1985). Mechanisms, e.g., ongoing staff development, use of evaluation, must be put in place to keep the innovation relevant and to prevent the change from giving way to business as usual.

(1). **Usefulness of the evaluation information to project participants.**

The findings of the case studies demonstrate that the staff within Title VII projects have been developing process evaluation activities to address their needs for information that will assist them in decision-making about the projects. Project staff referred to the need for timely information on the implementation of the projects and for information on outcomes other than student data on standardized tests. For these reasons, project staff frequently reported that the process evaluations they were carrying out were more useful to them than the summative evaluations, which arrived late and were focused primarily on student achievement test outcomes.

The approach being taken by many of the projects reflects recommendations found within the change process literature. For example, Guskey and Sparks (1991) argue that a multifaceted approach to evaluation is needed in order to produce meaningful and sustained change and improvement. They outline several guidelines for evaluation of improvement efforts, including: evaluation should begin during planning phases and continue throughout all phases of implementation; there should be a recognition that changes in one part of a system are likely to affect other parts as well; all parts of the educational system should be involved (i.e., curriculum, district and school leadership, parents, etc.); evaluation should address contextual factors related to the implementation of the improvement effort; evaluation information should be used to improve the program as well as to make judgments about it; and improvement efforts should be driven by clear objectives expressed in terms of student outcomes (Guskey & Sparks, 1991). In general, they note that a multifaceted approach to implementation of change is important just as teachers in classrooms recognize that a variety of strategies are more likely to be effective than the use of a single strategy for all students at all times (Guskey & Sparks, 1991). It appears that Title VII projects are attempting to take this type of evaluation approach, as seen in the frequency of process evaluation activities and the use of alternative means of assessing project outcomes.

(2). **Dissemination of evaluation findings.**

Evaluation findings cannot be helpful if they are not made available to the participants in a project. The case studies showed that

dissemination efforts varied a great deal. The dissemination of evaluation reports and discussion of findings in site #18, for example, included all participants in the project in briefings and discussions of the project's progress and outcomes, from district personnel through to teachers and parents. In site #15, dissemination included all participants and feedback was provided through midyear reports and brief reports at other times, as well as annual reports.

At the other extreme, however, as an example of poor dissemination, in site #12, only the district superintendent was provided with a copy of the evaluation report. In several of the other case study sites, teachers were given oral briefings on the evaluation findings, but it is not clear how thorough these were and, therefore, how well-informed teachers were regarding the evaluation outcomes. For example, in site #10 the master teacher was provided with evaluation findings. She was frustrated, however, because she wasn't given sufficient assistance in understanding the data to allow her to present it and explain it to the teachers on the project.

The research on change processes suggests that dissemination activities should place a special focus on presenting evaluation results to teachers. Change can be promoted through dissemination, since teachers are more likely to use those innovations that they perceive as working. In fact, teachers can become committed to use of an innovation through seeing positive outcomes as they implement the change. In addition, dissemination to those outside of the project can assist in the process of building support for the projects' activities; within the case study projects, this was mentioned as one use of the summative evaluations, for instance, to obtain support for the project by demonstrating the positive outcomes.

2. Direct concern for capacity-building: Evaluation of progress toward institutionalization

As a second function of evaluation in development of capacity building, evaluation can address progress toward institutionalization of the improvement directly through examination of the steps taken toward institutionalization and the level of success in achieving capacity-building goals.

Huberman & Crandall (1983) reported that achieving the mastery of complex changes was a process that took some 18 months and that after that period of initial implementation the steps taken next were critical to the future of the innovation. If the new practices are built into the regular training, staffing, budget cycles, then they survive; if not, they are gradually lost. Fullan (1985) suggests some steps that can be taken to prevent the loss of an innovation through attrition or neglect: develop plans

to train and assist new teachers, incorporate the new practice into formal curriculum plans and job descriptions, allocate regular budget line items for materials, and include orientation to the innovation and expectations regarding its continuation when replacing administrators or other leaders.

Despite the objective of Title VII funding to develop seed projects that will build new local capacity, only two case study sites (sites #13 and #16) specifically addressed capacity building within the evaluation. Site #13 referred to one purpose of the evaluation as being to demonstrate that the project would be continued by the district after the project ended. In site #16, capacity building was a specific outcome objective and was viewed from several perspectives: assumption of costs, development of district training plans and criteria to ensure continuation of the staff development efforts, and coordination of training activities with local colleges, universities, the MRC, and the State Department of Education. The lack of capacity-building as a specific objective in the other case study sites indicates that there is need for greater attention to this on the part of the Title VII projects.

Finally, an ultimate goal with regard to capacity-building would be to develop through the implementation of a particular change effort or efforts gains in "change process capacity" (Huberman & Crandall, 1983; Fullan, 1985). In other words, an additional project goal can be to develop the attitudes and skills that support innovation, such as those involving input and collaboration from teachers, restructuring of principal-teacher relationships, and leadership skills on the part of teachers as well as principals. In fact, through the process of implementing any improvement effort, it is advisable to monitor the process itself. As Fullan (1985) has pointed out: "Change involves pressure, assistance, and skills, but people must feel good about their relationships, sense of community, and sense of progress that result from their efforts". If the process of change itself is not monitored, a school or district could find that a highly successful improvement effort could demand so much from those involved that they are unwilling to attempt another significant change for a long time (Fullan, 1985).

## 3. Implications of the Change Process Literature for Title VII Evaluations

This brief overview and discussion of the literature on change and improvement processes has demonstrated that if the goal of local Title VII project improvement efforts is to develop new capacities within the school or district, then, in addition to the monitoring of student outcomes, other aspects of the improvement effort require attention. These are elements of the improvement effort such as:

- the actual process of implementation (e.g., who is involved in decision-making, the relationships among participants such as collegiality, information-sharing, etc.);

- the nature of the staff development process (e.g., nature and amount of assistance/follow through; development involving teachers as learners);

- the extent/quality of implementation of the improvement effort (e.g., to what extent are staff utilizing the new approach in the classroom, how well is the new approach/materials being implemented);

- the degree of support for the innovation (e.g., support for the innovation from participants and from administrators such as principals and district staff);

- the degree of "fit" within the school/district context (e.g., "fit" with local needs and "fit" or linkage with other services or program elements present in the school/district);

- the nature of information gathering/evaluation (e.g., usefulness of information to those involved in the change effort, both as a measure of the success of the innovation and as a guide to areas for further improvement); and,

- a direct concern for capacity-building (e.g., examination of specific actions taken to help ensure that the innovation becomes institutionalized and continues after funding).

Based on the literature on local change and improvement processes, it is clear that evaluation can serve an important function in effective implementation and in efforts to ensure the continuation of an innovation. Thus, process evaluation activities are central to the development of a sustained improvement a local school/district's capacity to provide effective services to limited English proficient students. In addition, evaluation activities that are directed toward an examination of capacity-building as a specific project outcome are important in order to assess the degree to which progress toward institutionalization is being made.

## D. Summary of the Case Study Findings

The case studies of the evaluations carried out in 18 Title VII projects revealed a pervading dissatisfaction with the Federal regulations' reliance on summative evaluation activities. Project staff were also unhappy with the focus on student achievement test data as the measure of project outcomes and felt there should be alternative means of examining the progress of the projects toward their stated objectives. As a consequence of this dissatisfaction, projects have turned to process evaluation activities to address their needs for information. Project staff reported that process evaluation findings were far more useful to them than the summative evaluation reports prepared for the Department of Education.

This movement in the direction of process evaluation by the Title VII projects, independent of the Department of Education requirements, is given strong support in the literature local change and improvement processes. The literature has demonstrated that process evaluation activities are essential to supporting local project efforts toward effective implementation and institutionalization of innovations. The implication of this finding is that the Department

of Education should support the projects in their efforts in carrying out process evaluation activities.

At present, there is no systematic approach in Title VII to process evaluation, and much of what the projects are doing is not fully consistent with the recommendations found in the literature. Therefore, the projects need guidelines and assistance in this area.

In addition, the case study findings indicated several weaknesses in the summative evaluations that were being carried out. The data elements required by the Title VII regulations are not consistently provided by the projects' evaluation reports and, in addition, not all data reported are valid and reliable. The projects appear to need further assistance in developing and working with the data required.

As a summary of the findings, the following recommendations are suggested by the case studies:

1. Include process evaluation within the Title VII evaluation requirements.

   The summative evaluations currently required focus on addressing the Department of Education's needs for an assessment of project effectiveness. However, these evaluations do not address the local project's needs for information. Inclusion of process evaluations to address the local project's needs and goals would also ultimately fulfill the Department of Education's goals for increased local capacity.

2. Provide assistance to projects in carrying out process evaluation activities.

   Since there has not been much documentation overall of the process evaluation activities within the project evaluation reports, it is not clear how well these are being carried out. While formal procedures are in place and activities are well documented in some projects, in others there is little or no documentation and the nature of the efforts are unclear; in still others, project staff express an interest in receiving assistance in carrying out process evaluations. Projects need guidelines and assistance in designing and carrying out process evaluation activities that effectively assess their specific project efforts.

3. Promote more ongoing assistance to projects in carrying out evaluation-related activities.

   The case studies have shown that problems arise in the course of the year which, when not addressed, result in poor data and weak evaluations. Increased contact of the evaluator with the project staff over the course of the year would help to identify problems early and resolve them before they seriously weaken the project and a project's evaluation. Within the 18 case study projects, many evaluators did not have much contact with the day-to-day operation of the project, primarily dealing with the project at the end of a project year; thus increasing an evaluator's contact with a

project over the course of the project year would provide needed additional assistance.

One final issue that has been identified in the course of the case study effort concerns the question of when OBEMLA should be notified of modifications made in project activities or evaluation plans, and the extent to which OBEMLA approval is required for any modifications. For example, in site #12 the project director felt that changes should be made in the project model and evaluation plan in order to better fit within the parameters of the overall district program for LEP students. However, the project director did not attempt to make such changes, explaining this by saying that since the project had received no feedback from OBEMLA on minor issues they had raised, they did not feel comfortable recommending major changes in the objectives. In district #3, on the other hand, due to various factors including a shift in the district approach to instruction of LEP students, the project shifted its goals and the focus of its activities. The evaluators regarded this shift as unfortunate and felt that the project had shifted too far from its original plan. Similar issues arose in other of the case study projects as well.

Projects therefore need guidance regarding the extent to which modifications can be made without approval when project staff determine these are needed, for example, modifications recommended based on findings from an examination of the implementation process or due to changes in local circumstances. In addition, if approval is required, projects need to know what is necessary in order to request approval for a change and what the process for this change should be. The projects will also need to have the confidence that such requests will be addressed in a timely manner, so that project implementation and planning is not delayed.

## References


Berman, P. & McLaughlin, M.W. (1978). Federal programs supporting educational change. Vol. 8. Implementing and sustaining innovations. Santa Monica, CA: Rand.

David, J.L. Purkey, S., & White, P. (1989). Restructuring in progress: Lessons from pioneering districts. Washington, D.C.: National Governor's Association.

Deal, T. E. (1990). Reframing reform. Educational Leadership, 47, 8, 6-12.

Elmore, R.F. & McLaughlin, M. (1988). Steady work: Policy, practice, and the reform of American education. Santa Monica, CA: RAND Corporation.

Fuhrman, S., Clune, W., Elmore, R. F. (1988). Research on education reform: Lessons on the implementation of policy. Teachers College Record, 90, 2, 237-257.

Fullan, M. (1985). Change processes and strategies at the local level. Elementary School Journal, 85, 3, 391-421.

Fullan, M., Bennett, B., & Rolheiser-Bennett, C. (1990). Linking classroom and school improvement. Educational Leadership, 47, 8, 13-19.

Glickman, C. (1991). Pretending not to know what we know. Educational Leadership, 48, 8, 4-10.

Goldman, C. & O'Shea, C. (1990). A culture for change. Educational Leadership, 47, 8, 41-43.

Guskey, T.R. (1985). Staff development and teacher change. Educational Leadership, 42, 7, 57-60.

Guskey, T.R. & Sparks, D. (1991). What to consider when evaluating staff development. Educational Leadership, 49, 3, 73-76.

Huberman, M. & Crandall, D. (1983). People, policies, and practices: Examining the chain of school improvement. Vol. 9. Implications for action, a study of dissemination efforts supporting school improvement. Andover, MA: The Network.

Huberman, A.M. & Miles, M.B. (1984). Innovation up close. New York: Plenum.

Lieberman, A. & Miller, L. (1990). Restructuring schools: What matters and what works. Phi Delta Kappan, 72 (June), 759-764.

McDonnell, L.M. (1985). Implementing low-cost school improvement strategies. The Elementary School Journal, 85, 3, 427-437.

McLaughlin, M. (1990). The Rand Change Agent Study revisited: Macro perspectives and micro realities. Educational Researcher, 19, 9, 11-16.

Miles, M.B. & Louis, K.S. (1990). Mustering the will and skill for change. Educational Leadership, 47, 8, 57-61.

Purkey, S. & Smith, M. (1985). School reform: The district policy implications of the effective schools literature. Elementary School Journal, 85, 3, 353-389.

Purkey, S. & Smith, M. (1983). Effective schoools: A review. Elementary School Journal, 83, 427-452.

Richardson, V. (1990). Significant and worthwhile change in teaching practice. Educational Researcher, 19, 7, 10-18.

Roberts-Gray, C. & Gray, T. (1983). Implementing innovations: A model to bridge the gap between diffusion and utilization. Knowledge: Creation, Diffusion, Utilization, 5, 2, 213-232.

Saphier, J. & King, M. (1985). Good seeds grow in strong cultures. Educational Leadership, 42, 6, 67-74.

Simpson, G. (1990). Keeping it alive: Elements of school culture that sustain innovation. Educational Leadership, 47, 8, 34-37.

Sparks, G., Nowakowski, M., Hall, B., Alec, R., & Imrick, J. (1985) School improvement through staff development. Educational Leadership, 42, 6, 59-61.

Stallings, J. & Mohlman, G. (1981). School policy, leadership style, teacher change, and student behavior in eight high schools. (Final report). Washington, DC: National Institute of Education.

Tillema, H. & Koster, D. (1990). Impact of educational support programs in schools. Knowledge: Creation, Diffusion, Utilization, 12, 1, 43-52.

Turnbull, B.J. (1985). Using governance and support systems to advance school improvement. The Elementary School Journal, 85, 3, 337-351.

Wood, F., Freeland, R., & Szabo, J. (1985). School improvement is more than school improvement. Educational Leadership, 42, 6, 63-66.

t7\secsumm(3)



ED/OPP93-32

*U.S. G.P.O.:1993-348-589:80203