DOCUMENT RESUME

ED 321 545 FL 018 631

AUTHOR	Lam, Tony C. N.; And Others
TITLE	Refinement and Field Test of Evaluation Procedures and Materials for ESEA, Title VII Bilingual Education Projects. Phase III Report.
INSTITUTION	RMC Research Corp., Mountain View, Calif.
SPONS AGENCY	Department of Education, Washington, DC.
PUB DATE	2 Jul 87
NOTE	142p.
PUB TYPE	Reports - Descriptive (141) -- Tests/Evaluation Instruments (160)
EDRS PRICE	MF01/PC06 Plus Postage.
DESCRIPTORS	*Bilingual Education Programs; Elementary Secondary Education; English (Second Language); *Evaluation Methods; Federal Programs; *Field Tests; *Program Evaluation; School Districts
IDENTIFIERS	*Bilingual Education Evaluation System

ABSTRACT

This report describes an extensive field test of the Bilingual Education Evaluation System (BEES) used to evaluate local level bilingual education projects. Because such projects will usually not be able to implement a traditional true or quasi-experimental design, BEES employs a "gap-reduction" evaluation design that is easily implemented, satisfies regulatory requirements, and yields an index relating to the growth made by project participants without requiring a comparison group of students. This evaluation method was tested at nine school districts in six states. This report describes this field test and the training and assistance program offered to participating educators. The report also summarizes the participants' comments and suggestions about how the evaluation system might be improved. In particular, it is suggested that the written materials be improved and that user-friendly software be provided to assist in the database management tasks and to perform the gap-reduction computations. Extensive documentation of the field test is appended including survey instruments and sample forms. (JL)

REFINEMENT AND FIELD TEST OF EVALUATION PROCEDURES AND MATERIALS FOR ESEA TITLE VII BILINGUAL EDUCATION PROJECTS

PHASE III REPORT

Tony C. N. Lam
Nona N. Gamel
G. Kasten Tallmadge

July 2, 1987

TABLE OF CONTENTS

LIST OF TABLES

I. INTRODUCTION

The Evaluation of Bilingual Education Programs for Language-Minority, Limited-English Proficient Students, currently being conducted by RMC Research Corp. under contract to the U.S. Office of Education, is intended to improve local evaluation practices with the dual goal of enhancing local utility of evaluation information and providing a data base that will be useful for broader purposes. As a first step in the study, an extensive review of the literature was conducted to determine the state of the art in bilingual education evaluation in the United States and to develop recommendations for a bilingual evaluation system. The second step was the development of a methodologically sound, standardized evaluation system and its documentation in a two-volume Users' Guide. The Bilingual Education Evaluation System (BEES) was designed for use in evaluating local level bilingual education projects.

Development of the BEES was guided by four primary design objectives:

- The system should reflect the knowledge gathered out of previous work in bilingual education evaluation.
- It should provide guidance on how to avert salient threats to the validity of evaluation findings.
- It should be useful at the local level for purposes of project improvement.
- It should be totally responsive to the current federal legislation and regulations governing the evaluation of Title VII projects.

In developing the BEES we assumed that most projects will not be able to implement a traditional true or quasi-experimental design for reasons that are well documented throughout the literature. For this reason, we formulated a "gap-reduction" design that is easy to implement, satisfies

the regulations' requirements, and does not require a comparison group made up of students similar to those served by the project.

The gap-reduction design yields an index relating the growth made by project participants to that made by a non-project comparison group rather than a quantitative estimate of the project's impact. The evaluator must rely on a wide array of information and findings about both process and outcome to form judgements about project effectiveness. The system emphasizes quality control and recognition of unavoidable problems as a method of avoiding erroneous conclusions.

The third step of the study was to field test the system in nine school districts across the country. The purposes of the field test are to answer the following questions about the BEES:

- To what extent has the system aided the sites in defining and objectively measuring the impact of the bilingual program on its students?

- To what extent can the system serve to improve local evaluation practices?

- To what extent can the system generate data suitable for use in local decision making?

- How adequately can the evaluation designs be integrated with a district's regular instructional and testing programs?

- How suitable are data generated by the system for synthesis across Title VII projects?

- What are the costs incurred in implementing the system's evaluation designs (costs associated specifically with the project site, start-up costs, and ongoing costs)?

- To what extent would a district require technical assistance in implementing the system?

While participation in the field test has represented some burden to the sites, this has been offset by training and technical assistance provided by RMC staff. These activities will be described in the next chapter of the report.

Our goal in selecting field test sites was to represent the broadest possible range of bilingual education projects in terms of geography, minority languages, and program design. Site selection criteria are described in Appendix A. We decided that because California, Texas, and New York have the largest bilingual populations, they should each have two sites. Arizona, Florida, and Oregon were selected to participate with one site each.

SEA's were asked to nominate sites for inclusion in the study, and sites were then asked to participate. Appendix B is the invitation letter sent to sites. Two initially selected sites dropped out and were replaced. We were not able to locate any late-exit or immersion projects willing to serve as field test sites. Each field test site is described briefly below and summarized in Table 1.

Tucson Unified S.D., Arizona

The Tucson Unified School District serves 54,000 students, of whom 26% are language minority and 22% are Spanish speaking. Language-minority students are over-represented in special education (13.5% compared to 8% district-wide) and under-represented in gifted classrooms (.3% compared to 4% district-wide). Even lower percentages of LEP students are in gifted classrooms.

Bilingual special education classes are funded by the district. The Title VII project supplements and expands these programs in grades K-12

and has enabled the establishment of a bilingual gifted class for gifted LEP students. The project offers transitional bilingual education to identified students. When the students become English proficient, they have the option of transferring to comparable all-English classrooms or of remaining in bilingual education. Approximately 228 students are served, of whom 85% are low income.

San Jose Unified School District, CA

San Jose Unified School District, located south of San Francisco, serves approximately 217,704 students (of whom 85% are from low-income families and 12% are LEP) representing at least 36 different languages. One of the fastest growing language groups in the district is Indochinese. The Indochinese LEP population grew from 455 students in 1981 to 766 in 1986.

The Title VII project is located in five schools, elementary through high school, which were identified as having a significant influx of Indochinese students and having demonstrated a long-term commitment to meeting the needs of Indochinese LEP students. The project includes ESL instruction for all Indochinese students, L1 instruction for Vietnamese-speaking students and for other language groups where possible, staff and parent training, and curriculum development.

Valley Center Union School District, CA

The Valley Center Union School District, in a rural part of San Diego County, serves over 1,300 students at grades K-8. The population includes commuters, ranchers, citrus farmers, and ranch- and farmworkers. Approximately 12% of the district's students are Spanish-dominant LEP students. The district has provided ESL and Spanish reading instruction to these students, who by grade 8 lag approximately five years below the grade level of their English-speaking peers while being chronologically two years older.

The project is a transitional bilingual education project emphasizing concept and skill development in Spanish with later transfer to English. The project will serve 162 LEP students at grades K-8, 74% of whom are from low-income families. The project uses as basal readers language texts centered around each child's unique experiences. The project also includes Spanish language development for English-only students.

Community School District #5, New York City

Community School District #5, in economically-deprived Harlem, has approximately 11,911 students enrolled in its public schools and 3,162 in non-public schools. In the public schools, there are 836 LEP Hispanic students. The families of many of these children have recently immigrated from rural areas offering few or no school services. Over half the parents have less than an eighth-grade education.

The Title VII project is located in P.S. 161, an elementary school with a large concentration (over 80%) of Hispanic students which continually receives new arrivals from Hispanic countries. This school also houses a district bilingual program and a Chapter 1 program. The Title VII project will serve new arrivals in grades K to 3, new incoming kindergarten and first graders, and second and third graders who score at or below the 20th percentile on the LAB, a total of about 258 students. The project is a transitional bilingual education project with an emphasis on developing students' skills in both Spanish and English.

School Board of Polk County, FL

Polk County is a 58,000 student school district located in central Florida. Citrus cultivation, one of the county's main industries, attracts a large population of Spanish-speaking migrant farmworkers to the county. In 1981-82, Polk County schools served 6,683 children of migrant workers. Four percent of the district's minority-language students are referred each year for serious intellectual, academic, social, and/or emotional problems.

A district-funded bilingual education program has been in place since 1979-80. The Title VII transitional bilingual education project has been implemented to bring additional services to district schools having the highest concentrations of LEP Hispanic students. Goals of the project are to meet the academic needs of these LEP students to enable them to achieve at the rate and level of their non-LEP peers and to provide students and their families with needed services through employment of a Hispanic social worker. The project serves approximately 380 students in grades K-8.

Tyler Independent School District, TX

Tyler Independent School District in Northeast Texas is a district of 16,500 students, of whom 3% are Spanish-dominant and LEP. The number of LEP students has been increasing dramatically. These students on the average are achieving from one to three years below grade level, and most are from families of very low socio-economic status. The district suffers from a shortage of teachers with bilingual education credentials.

The Title VII transitional bilingual education program, which serves approximately 300 students, operates out of two elementary school buildings in attendance areas having the largest numbers of LEP students. Students in other attendance areas are bused to one of these buildings. Staff development, in conjunction with a nearby university, is an important part of the project, so that local Spanish-speaking teachers and prospective teachers can obtain appropriate credentials.

Valley View Independent School District, TX

Valley View Independent School District, located in a rural part of Texas along the Rio Grande, serves 746 students. Of these 98% come from low-income families and 88% are LEP. Ninety-five percent of the district's population is of Mexican descent, and 65% are migrant field workers, many of whom are recent immigrants. LEP third and fifth graders in

the district achieve less than 50% mastery of objectives on the Texas Assessment of Basic Skills.

The Title VII project is located in the district's elementary school site. It serves approximately 550 students in grades K-5 with a program emphasizing concurrent development of English and Spanish reading skills. The project has a materials development component involving the development of computer software for the reading instruction. Forty parents of children with severely limited English proficiency will be recruited for a special parent involvement program emphasizing literacy and transfer of skills to English.

Salem-Keizer Public Schools, OR

In the fall of 1985, Salem-Keizer Public Schools began serving district five-year olds for the first time. Half of district students and over 80% of district LEP students had in the past been found unable to enter first grade at age six. Because the district offers bilingual programs at first and successive grades, there was a need to offer a bilingual program for five-year old LEP students who would then continue with a bilingual program in the first grade.

Four elementary schools with the highest concentrations of LEP students in the district were selected for the program, which includes students whose primary languages are English, Vietnamese, Khmer, and Spanish. The goal of the program is to enable studerts to be promoted to first grade and through successive grades. The program includes a home-teaching component for parents, and serves approximately 100 students.

Community School District #10, New York

Community School District #10 is located in the northwestern portion of the Bronx. It has a highly diverse, multi-ethnic student population of about 35,000 students in grades K-9. Approximately 20% are LEP students

of Hispanic or Asian backgrounds. Four intermediate and junior high schools in the district have been identified as most in need of assistance, based on mobility and truancy rates and percentages of LEP students. Over half of these LEP students are Hispanic.

The Title VII project provides supplementary services and opportunities to those Spanish-dominant students judged most in need of assistance. Services include ESL; self-awareness; decision making; career, economic, and educational awareness, and employability skills. Students also receive assistance from a bilingual education assistant and a family worker. Approximately 300 students are served.

Contents of the Report

This report describes the field test of the bilingual evaluation system in these nine districts. Chapter 2 discusses training and technical assistance offered by RMC to the field test sites to assist their implementation of the evaluation system. Chapter 3 summarizes comments and suggestions for changes for each of the 10 chapters in Volume 1 of the Users' Guide gathered during the field test, and Chapter 4 summarizes the lessons learned from the Field Test.

Table 1

Site	Current Project Year	No. Students Served	Grade Levels Served	Predom. Language Groups	Program Type
Tucson Unified S.D., Arizona	3	228	K-12	Spanish	TBE-special ed. and gifted classes.
School Board of Polk County, FL	1	380	K-8	Spanish	TBE and support services for migrants.
Community S.D. #10, New York	1	300	6-9	Spanish	ESL, career education
San Jose Unified School District	3	360	K-12	Vietnamese, other Asian	ESL and TBE
Community School District #5, New York	1	258	K-3	Spanish	TBE, emphasis on thinking skills
Valley Center Union S.D., CA	3	162	K-8	Spanish	TBE, home teaching component
Salem-Keizer Public Schools, OR	2	99	age 5	Multiple	TBE
Tyler Independent S.D., Tyler, TX	1	300	K-5	Spanish	TBE, emphasis on staff development
Valley View Independent S.D., TX	3	550	K-5	Spanish	TBE, materials development, parent education

II. SRA TRAINING AND TECHNICAL ASSISTANCE ACTIVITIES

Two workshops were scheduled during the field testing, one in October, 1986 and one in June, 1987. The purpose of the first workshop as to train the field test participants on the bilingual Education Evaluation System (BEES). The June workshop or exit conference was designed to provide an opportunity for these participants to share their experience in trying out the BEES and to provide input for further revision of the Users' Guide. Agendas and correspondence relating to these workshops can be found in Appendix C.

The October workshop was held from October 7-9 in San Francisco. During the first morning, Lam presented an overview of the BEES and discussed evaluation planning and focusing in the BEES. In the afternoon, Lam discussed process evaluation. Workshop participants took a pretest at the beginning of the morning session, and a posttest at the end of the day so that they could measure how much they had learned.

On the second day, Tallmadge discussed outcome evaluation, including measurement error and the regression effect, selecting achievement tests, other indicators of achievement, test administration and scoring, calculating growth indices, and quasi-experimental designs. Workshop participants took a pretest designed to sensitize them to the issues that would be presented that day. There was no posttest, however.

On the third day, Tallmadge and Lam discussed data processing and analysis, integration and interpretation of results, and evaluation reporting. At the end of these presentations, the workshop turned to an open discussion of the Users' Guide and the field test.

After this meeting, follow-up letters were sent to participants to clarify their roles in field testing the Users' Guide. This letter highlighted key activities desired of field test site staff, asking them specifically to implement Users' Guide recommendations or document why the

recommendations were not practical in their districts. A similar letter was sent to the sites in April in preparation for the June, 1987 workshop.

The June workshop was held on June 17-19 in Washington, D.C. This workshop was attended by four representatives from Tyler, Texas; two representatives each from Community School District #5, New York; Community School District #10, New York; Valley Center, California; San Jose, California; and Valley View, Texas; and three representatives each from Tucson, Arizona; Salem-Keizer, Oregon; and Polk County, Florida. Additional participants during all or part of the workshop were James English and Keith Baker from the Education Department, Office of Planning, Budget, and Evaluation; Pat Johanson, from the Education Department, Office of Bilingual Education and Minority Language Affairs; and various staff from the Education Assistance Center - East. Tallmadge and Lam represented the project staff. The first day of the workshop was devoted to presentations by each of the field-test sites. A staff member from each site presented an overview of the site's Title VII project, the activities undertaken to implement each Users' Guide recommendation, and a description of the results of the implementation, including any problems.

The second day and part of the third morning was devoted to a chapter-by-chapter review of the Users' Guide by field site staff. Site project staff divided into four subgroups (New York districts; Salem and Tucson; Valley Center, San Jose, and Valley View; and Tyler and Polk County). Tallmadge met separately with attending project evaluators to review Chapters 6 (Implementing the Gap-Reduction Design) and 8 (Processing and Analyzing Data). Review of the chapters by project staff took longer than expected and, as a result, only the first seven chapters were covered during the workshop (Chapter 8 was reviewed by evaluators only).

Two questionnaires were also administered during the workshop (see Appendix D). One questionnaire was designed to ascertain the participants' perception of the importance of the evaluation data recommended for collection on students, instruction, and staff for estimating program

effectiveness and for improving project performance. This questionnaire was a follow-up of another questionnaire sent to the participants prior to the June workshop. In this earlier questionnaire, each pilot site was asked to rate the availability and accessibility of the data included in the later questionnaire. The second questionnaire that was administered during the workshop was designed to obtain input from the nine sites regarding the importance and usefulness of the community characteristics, district demographics, and school characteristics data recommended by the BEES. The results of these three questionnaires are discussed in Chapter 3.

Part of the third morning was used to summarize findings from the first two days. The findings from this meeting and from the field testing with regard to the implementation of the Users' Guide are presented in Chapter 4.

In accordance with the Phase III modified contract, two site visits were made to each site by either Lam or Tallmadge. On the average, each visit lasted two days. Table 2 lists the dates for all these site visits. As Table 2 shows, site visits were scheduled at varying intervals, depending upon the availability of RMC staff and the wishes of site personnel. Lam visited Tucson, San Jose, Valley Center, New York Districts #5 and 10, and Salem. Tallmadge visited Polk County, Florida, and Tyler and Valley View, Texas.

The conduct of each site visit, was quite similar. Each visit included a meeting with the site's bilingual project staff, individually or in a group, to discuss the site's program and evaluation designs as they related to the recommendations embedded in the Guide. For example, project staff might have been concerned that the test used by the project to measure academic achievement was too difficult for the target population. Since the Guide recommends the use of functional-level testing, this alternative testing procedure would have been proposed by RMC staff and discussed with the group.

In addition, for some sites, we conducted classroom observation with the project site staff either to determine the project's critical instructional features or to try out a newly developed classroom observation checklist. The sites for which classroom visitation was conducted include San Jose, Valley Center, Salem, and Tucson.

In addition to site visits and these workshops, we offered sites technical assistance through frequent telephone conversations and periodic correspondence.

Since outcome data were not available, the technical assistance provided thus far has focused primarily on evaluability assurance, evaluation planning, process documentation, data collection and management, and assessment instrument review. It is anticipated that after the June meeting, limited technical assistance will be provided on data analysis and reporting. In Table 3, different categories of technical assistance activities and the sites receiving each type are summarized.

Table 2 - Site Visitation Dates

Site	1st Visit	2nd Visit
Tucson Unified School District, AZ	1/7-8/87	5/11-13/87
San Jose Unified School District, CA	10/24/86 1/22/87	5/7/87 6/9/87
Valley Center Union School District, CA	2/16-17/87	4/27-28/87
Polk County School District, FL	3/17-18/87	5/12-13/87
Community School District #5, NY	3/18/87 3/20/87	5/27-28/87
Community School District #10, Bronx	3/17/87 3/20/87	5/26-27/87
Salem-Keizer Public Schools, OR	2/23-24/87	5/4-5/87
Tyler Independent School District, TX	3/10-11/87	5/14-15/87
Valley View Independent School District, TX	4/7-8/87	6/5/87

Table 3. Areas of Technical Assistance by Site

Areas of Technical Assistance	Site								
	Tucson	Salem	Polk County	Tyler	Valley View	Valley Center	San Jose	NY #10	NY #5
o Revise project objectives	X	X				X	X	X	X
o Elaborate on project design	X	X				X	X	X	X
o Develop classroom observation checklist	X	X	X	X		X	X	X	X
o Review data to be collected	X	X	X	X	X	X	X	X	X
o Manage data	X	X		X	X	X	X		X
o Review tests and testing practices	X	X	X	X	X	X		X	X
o Review affective outcome instruments and methods of effective assessment	X					X		X	
o Data analysis (including gap-reduction analysis)			X	X	X				
o Reporting									

In the remainder of the chapter, we describe, site-by-site, our consultation with the nine field-test sites. We focus on the unique evaluation problems encountered by each site, the recommendations provided by us in dealing with these problems based on the Guide, the problems in implementing these recommended practices, and the sites' needs for additional technical assistance.

Tucson Unified S.D., Arizona

The objectives stated in the Tucson Unified School District Title VII project's proposal are all process objectives that deal with identifying and assessing LEP gifted students, establishing a bilingual gifted class, expanding bilingual special education, training staff and involving parents. Expected student outcomes are only briefly mentioned.

As stated in the proposal, students' literary skills in Spanish and English, and their math skills are expected to be enhanced as a consequence of their participation in the project. Spanish reading skills and math skills are measured by the Bateria Woodcock (K-12), La Prueba (K-8), and the Brigance Diagnostic Assessment of Basic Skills (K-8) tests. Skills in English reading are measured by the Woodcock and the Brigance tests after the students are exited from the bilingual to the regular special-education classes.

The La Prueba test is a norm-referenced test and is administered to all district LEP students. The Woodcock test is a norm-referenced test and the Brigance test is a criterion-referenced test. Both tests are used for identifying students in need of special educational service and for diagnosing the students' strengths and weaknesses. This diagnosis forms the basis for developing an Individualized Education Plan (IEP) for each student. This individualized instructional approach has made it difficult to establish a general criterion of success (as measured by test scores) for all students, since the expected growth rate varies among students.

One recommendation for assessing the project's effect on student outcomes is to examine each student's performance individually. One way to accomplish this is to calculate the percentage of incorrect responses on the Brigance pretest (each student is administered a different set of items depending on his functional level) that are correctly answered during posttest. Since each student has his own set of objectives, another way to assess project impact is to compute the percentage of objectives completed, partially completed, and not completed for each student. Criteria for success for all outcome objectives based on these statistics, and project performance can be evaluated based on the extent to which these criteria are met. Undoubtedly, stakeholder bias could be severe with this approach.

The La Prueba test results will also be useful for comparing project and district LEP students. Since the test is designed for regular students, and not for students with physical, language and/or learning disabilities, functional-level testing should be used and some modification of the test instructions should be made. The project has done both. The test modifications include allowing students to mark their answers on the test booklet, extending the time limits, and giving more elaborate instructions. Since the test can only be used at grades K-8, the Woodcock test will have to be administered to students in grades 9 through 12.

Due to its individualized instructional approach, it is also difficult for the project to prescribe the specific teaching approach that all project teaching staff should follow. During the two site visits, efforts were made by Lam to assist the project staff in generating a list of critical instructional features. After much discussion and some classroom observation, it was decided that the observational scheme should incorporate the variety of the IEPs. Instead of specifying desired teaching behaviors, three general questions were developed to determine levels of program implementation through classroom observation. These questions are:

- Are the IEP objectives appropriate for the student?
- Is this instructional approach appropriate to the IEP objectives?
- Is the language appropriate for the student as determined by the language of instruction prescribed in the IEP?

Specific indicators to guide responses to these three questions will be developed after some initial classroom observations. In addition, a teacher survey was developed to assess primarily how teachers use the two languages to conduct instruction (see Appendix E).

Technical assistance was also provided in data management. In part because of the reluctance of the project teachers to test students and record test results systematically, test scores for the first two years of the project are so incomplete that they cannot be used for program evaluation. There was no system for filing evaluation data once they were collected. Since the district has decided to participate in the field test, efforts have been made to centralize the data and to ensure complete testing. Teachers have received in-service training on better data recording. We will continue working with project staff to develop a data management system.

Salem-Keizer Public Schools

As stated in the district's proposal, the project has four goals, related to establishing and implementing an instructional program for 5-year old LEP students, involving parents, and training project staff. The two student outcome objectives specify expected achievement in English competence as measured by the LAS, and the extent to which the project students will meet the requirements of the district's five-year-old program as measured by the Salem Skills Checklist developed by the district.

One of the major difficulties in conducting evaluation for the Salem project is testing 5-year olds. Standardized testing instruments are not readily available, which makes implementation of the gap-reduction design

difficult because of the lack of normative data. It is not clear if the LAS is appropriate for measuring achievement in English language proficiency for this age group. The problem of curriculum-test match may still exist for this grade level. Although it is subjective, the Salem Skills Checklist, which is designed to measure skill levels in physical coordination, perceptual memory, language, and mathematics appears to be a viable instrument for measuring student achievement. Since pretest results from the checklist are used to prescribe instructional activities and the students are posttested on those areas in which they receive instruction, there is a perfect curriculum-test match. Again, using the checklist makes it difficult to implement the gap-reduction design.

There is a third implicit student outcome objective which relates to increasing student understanding of language and cultural heritage. The objective is very difficult, if not impossible to measure directly. Other indirect measures of the attainment of the objective, such as frequency of cultural activities and student behavior and participation in these activities, were proposed.

Procedures to document the level of instructional service delivery consumed a great deal of effort. During the initial meeting with the project staff, we brainstormed on the critical instructional features to determine ways to document the extent to which they are carried out in the classes. The first classroom observation instrument developed was a behavioral checklist that requires tallying occurrences of specified behaviors. When Lam and the project staff tried it out in the classroom, they found that it was not feasible to record the many instructional and student activities that are going on in the class. It was decided that the delayed-report approach is more appropriate.

After a few revisions, an observation form was developed that the project staff feel is useful and appropriate for documenting the project's instructional service delivery (see Appendix F). The form can be used to record activities for several small groups or for a single group across

different time segments. The observation system requires the observer to write down for each group or each time segment, the instructional methodology employed and the language used for instruction and their appropriateness for the students involved. Although the form is easy to use, the observers must be very knowledgeable about bilingual and ESL pedagogical methods. This is the tradeoff between structured versus non-structured observation that needs to be addressed in the Guide.

Another area in which the project needs technical assistance is in data management. Evaluation data are being collected from different sources: forms, student records, and district evaluation offices' computer files. The project plans to utilize its IBM computer to manage all the data. Continued assistance to integrate all these data into a computerized data base will be needed in the coming academic year.

Polk County, Florida

Technical assistance/field test visits were made to the Polk County school district on March 17 and 18 and May 12 and 13, 1987.

The March visit. During the March meeting, activity was focused on a page-by-page review of the Users' Guide with the project director and the district evaluator. Those persons had studied the Guide and had assessed the feasibility of implementing its recommendations. Many of the recommended practices and procedures were already in place. No recommendations that were not already in place had actually been implemented in the district, but many were scheduled for implementation next year.

In some cases, Polk County felt that their current practices had at least some features that were superior to the Users' Guide recommendations. In other instances, they preferred our recommendations to their current practice. In a few instances, the site persons identified conflicts between official state or district policies and our recommendations. All of these specific comments are described below.

The only general comment that site personnel made was that the Users' Guide was calling for the collection and summarization of substantially more information than they felt was required by the regulations and more than they or anyone else could use productively. They pointed out that, even though the information called for was available, the personnel and dollar costs associated with assembling it were substantial. They therefore preferred to avoid assembling non-required information for which they could see no local utility.

Specific Comments.

- Table 1 (p. II-5) is confusing.

- Probably teachers and certainly aides should be excluded from reviewing draft reports (last bullet, Table 6, p. III-7). It is difficult enough to get reports out in a timely manner without this type of input.

- Parents and community members should certainly not be involved in conducting native language testing. They are not trained, not unbiased, and often not literate (3rd bullet, Table 6, p. III-7).

- School district administrators should not review draft reports (last bullet, Table 6, p. III-8).

- Emphasize developing an action plan for implementing recommended project improvements. Without an action plan, recommendations are likely to die (last bullet, p. III-10).

- Compiling all of the data called for in Table 7 (p. IV-3) is excessive. "Tell us what we have to collect." Note also that some of the items can be derived from others and need not be separately compiled.

- Text (from bottom of p. IV-2 to top of p. IV-8) is very difficult to follow because of the insertion of three lengthy tables. Note too that the only reference to Table 9 occurs in the middle of Table 8.

- Table 8 should include parent/community involvement activities and curriculum development activities. Note last sentence under Project Characteristics on p. IV-2.

- The information called for in Table 8 (P. IV-6) that derives from requirements for grant applications is excessive. It would have to be collected each year as it becomes obsolete rapidly.

- Recommendations to maintain close contact with project staff (last sentence of paragraph in middle of p. IV-8) should be strongly emphasized.

- Figure 2 on p. IV-10 (really a Table) is useless.

- Regarding the first Self-Report example in Table 10 (p. IV-12) --this is a cumbersome task -- should get this information from an interview.

- Table 11 (p. IV-13) refers to "specified behaviors." Clarify where these behaviors come from.

- Should note that behavioral checklist deals with individual behavior elements while coded behavior record deals with sequences of behaviors.

- Regarding recommendation that teachers not be informed as to when classroom observations will occur -- union agreement in Florida requires substantial advance notice.

- Pages V-1 and V-2 talk about language proficiency without defining it. Are we talking about oral language comprehension, oral language production, reading, writing, grammar, spelling, or what? Please provide some guidelines!

- Throughout Chapter VI, you should refer to means and medians, not just means. You told us earlier that medians would be preferable under certain conditions.

- Define scale score.

- Explain how to plot gap-reduction data a la Figure 3 (p. VI-3).

- Florida adjusts for the regression-effect bias by multiplying the gain by the test-retest reliability of the pretest. Will this work in the gap-reduction design?

- The first sentence of the middle paragraph on p. VI-10 is confusing.

- First bullet under Planning the Testing Schedule -- what kind of exceptions are possible? Suppose a district wants to implement a spring-to-spring testing schedule. When should they pretest entering kindergartners? Is fall pretesting OK? Or does fall pretesting force the district into a fall-to-fall schedule from all grade levels? Can the kindergarten pretesting be skipped altogether? This is a problem with the regulations, not the Users' Guide, but please provide some guidance.

- On p. VIII-3 you talk as if each school will have its own computer and do its own analyses. Do you not mean something else?

- On page VIII-6, first paragraph under Other Summative Outcome Analyses -- perhaps dropouts need to be sorted by type, such as moved to another school in the district, migrated seasonally and are expected to return, etc. for elementary grades there are very few dropouts in the sense that we usually think of dropouts.

- The first sentence in the next paragraph needs to be explained.

- On page VIII-7, second paragraph from the bottom, you are asking for too much -- especially since you acknowledge that none of this information is required by the regulations.

- Page VIII-8, exposure-to-treatment scores needs to be clarified.

- The stratified analyses on p. VIII-10 are not required by the regs. You should acknowledge that fact. (The top of p. VIII-11 presents that acknowledgement, but it should precede, not follow, the descriptions of the analyses).

- Include teacher certification under implementation data (second bullet on p. IX-1). Are "waivered" teachers working toward certification, etc.? Inclusion of a cultural component may also be important.

There were no comments on Chapter X, and the site visitor discouraged extensive comments on Chapter IX because a major restructuring of that chapter was already underway.

The May Visit. The May visit was substantially different from that conducted in March. The major focus was on technical issues associated with implementing the gap-reduction design. Considerable time was spent reviewing the conditions under which it is necessary to adjust for the regression-effect bias and the statistical procedures involved in making that adjustment. The Polk County evaluation staff is also responsible for

Chapter 1 evaluation and is familiar with the regression-effect adjustment currently required by the Florida SEA. Unfortunately, the Florida procedure is different from both the TIERS-recommended procedure and that recommended for the BEES. It is also basically flawed in that it applies a multiplication adjustment to the gain rather than an additive adjustment to the selection/pretest mean score.*

We reviewed the reason that the mean selection/pretest score is negatively biased, the conditions under which an adjustment is necessary, the factors that affect the magnitude of the bias, and the procedures that are recommended for removing it. We also discussed ways in which the gap-reduction model could be applied to non-test data (e.g. attendance rate). Because of the technical sophistication of the Polk County evaluator, some of our discussions were more in-depth than would have been appropriate elsewhere.

Tyler, Texas

Tyler was visited twice -- the first time on March 10 and 11 and the second on May 14 and 15, 1987. Tyler is a unique site in that it was just beginning the first year of its grant at the time the field test began. For that reason, it attempted to implement nearly all of our recommendations as an integral part of its start-up activities. By the time of our first visit, the site had actually tried out most of the process-related ideas in the Users' Guide.

*The impact of using the same set of scores for both selection and pretest purposes is that the pretest mean is spuriously low and the selected students will score higher on the next testing, even if there is no intervention. The problem with attempting to correct for this bias by applying a multiplicative adjustment can be easily recognized if one considers the situation in which there is no pre-adjustment gain. With zero gain it follows that the adjustment must also be zero. On the other hand, the pretest score will be too low and should be adjusted upwards (thus creating a negative gain).

March Visit. During the first visit we reviewed the Users' Guide page-by-page. Site personnel found that most of our recommendations could be implemented and said that they had found them helpful. Some of the recommended practices they had planned to implement before they became involved in the field test, but there were others they could not have thought of on their own but did find helpful.

A few of our recommendations conflicted with Texas law (e.g. modifying and/or translating tests and testing out of level). We talked extensively during both visits (as well as at the Exit Conference) about this conflict without finding a satisfactory resolution for it.

In-level testing is only a problem when students encounter the test's floor at pretest time. When that happens, however, "real" student gains will be underestimated by the gap-reduction design (and other evaluation designs as well) with the final result that the project will be made to appear less effective than it really is.

The most unfortunate consequence of the requirement to test in level is that the greatest impact will be on projects serving the neediest students. The achievement growth of students participating in such projects will be the most seriously underestimated, and truly effective projects may even appear to have negative impacts.

Specific comments follow.

Specific Comments

- 2nd paragraph on page I-3 talks about what the gap-reduction design does not yield. Should say what it does yield.

- 3rd paragraph on page II-1 -- Tyler has also involved parents in these preparatory activities.

- Table 1 on page II-5 is confusing and the examples are not helpful.

- Table 2 (p. III-2) contains too many activities in a difficult-to-follow sequence. Some grouping arrangement would be helpful.

- There is no need to describe the project director as possibly "unwilling or unable to direct the evaluation in detail." The early formation of a strong evaluation team is equally important even if the director is both willing and able.

- The recommendations to involve aides in evaluation activities is a good one -- especially if the aides are teacher trainees. But neither they, nor parents and community members, nor district administrators should be involved in reviewing the draft evaluation report (Table 6, pp. III-7 and III-8).

- The early pages of Chapter IV are poorly and inconsistently formatted, titled, and punctuated.

- The recommendation to keep student data by student is good. Perhaps it should be emphasized more strongly.

- The bottom paragraph of p. IV-2 should be moved to the top of p. IV-8.

- Too much documentation of pedagogical methods, materials, and techniques are called for. Could this be limited to critical features?

- Educational and professional qualifications should be documented for all teachers and aides not just summarized across the project. This is very important! (Table 8, p. IV-6).

- The information called for on parent advisory counsels and staff training activities is appropriate. We collect it and find it useful. (Table 9, p. IV-6)

- Table 9 (p. IV-7) is not referenced in the text.

- Figure 2 on p. IV-10 is not a data extraction form, it is a sample of extracted (and useless) data.

- Tyler uses all three types of self-report measures and finds them useful (p. IV-10).

- The recommendations on classroom observations (p. IV-11) are too non-directive. Tyler wants more specific guidance on what they should (or are required to) do.

- With regard to delayed report observations, how long a delay is recommended, permissible? (p. IV-11)

- Table 11 (p. IV-13) should specify behaviors - who "specifies" them and how? etc.

- Step 4 (P. IV-15) is helpful specific guidance for classroom observation, but information should be expanded and presented earlier (see 2 bullets above).

- Page V-7. Tyler cannot make any modification to standardized achievement tests (although they can modify instructions). Extending time limits, simplifying language, and translating tests are prohibited by Texas regulations.

- Tyler collected all of the student characteristics data called for in Table 13 because of our recommendations. Several suggestions were made regarding how redundancy could be reduced. (page VII-7)

- Use of raw scores is discouraged (Top paragraph on p. VIII-2). The reasons are not made clear. What is gained by making "difficult" score conversions?

- Tyler needs help with data base management. Users' Guide recommendations are inadequate. Could we recommend specific software or source of technical assistance? (p. VIII-2, bottom two paragraphs).

- Page VIII-3 implies that analyses are done at the school level. All of Tyler's analyses are done in the district office by district staff. Is this contrary to the Users' Guide recommendations?

- The Users' Guide says that "other summative analyses" need be done only for project participants (p. VIII-6, near top of page). The recommendation is unclear. What about former participants? What about comparison groups?

- What do you mean by a "trend analyses?" (last line on p. VIII-6). Please clarify.

- Please emphasize more strongly that transitional bilingual programs cannot be adequately evaluated on a one-year basis. A longitudinal approach is needed (p. VIII-13, third paragraph).

- The next paragraph (p. VIII-14, fourth paragraph) should include an example.

Tyler had no comments on Chapter X. Comments on Chapter IX were discouraged since a major revision of that chapter is currently underway.

May Visit. At the time of the May visit, Tyler was beginning to think about the outcome component of their evaluation and requested tech-

nical assistance on various aspects of the gap-reduction design. Of primary interest was the fact that the design could be applied to non-test data. We worked through examples using classroom grades (useful particularly for assessing the performance of reclassified LEPS) and attendance data.

Site personnel were shown that a simpler form of the design (one that does not require calculating RGI's or even standardizing growth estimates) can be useful for local-level project improvement. Since these topics are covered only superficially in the Users' Guide it was agreed that they should be explained in greater detail when that document is next revised.

We also discussed the origin and cause of the regression-effect bias and how to correct for it. Persons with evaluation responsibilities for programs other than Title VII attended parts of these technical assistance sessions.

The topics of functional-level testing and modifying tests to "enhance their psychometric integrity" were also discussed, but little progress was made toward resolving the conflict between Users' Guide recommendations and Texas regulations.

Valley View, TX. The Valley View school district in Pharr, Texas, was visited on April 7 and 8, and again on June 5th. Unfortunately, between the initial meeting of the field-test sites in San Francisco during the first week of October, 1986, and the first site visit, major administrative changes had occurred within the district. Both the District Superintendent and the Coordinator of Federal Programs were new. The external evaluator, with whom we had expected to work closely, had also been replaced. Not surprisingly, the turmoil caused by these major changes resulted in field-test activities being given a low priority. Still, the site was anxious to collaborate and to do what it could to assist us in meeting our objectives.

Valley View was unique among our sites for other reasons as well. Some 95% of its students are Hispanic LEPs and approximately the same percentage are poor (participating in the free-lunch program). All students in the district receive bilingual education services except for about 5% -- those students whose parents want them to receive no instruction in their native language. There is no mainstream in the sense that that term usually connotes.

In addition to those demographic characteristics already mentioned, Valley View is a very poor school district (the local tax base is minimal) that serves a rural and highly mobile population. Most of its students do not pursue their education beyond the ninth grade, and there is no high school in the district.

The district does not use a test-scoring service -- teachers score the tests. It has a non-functional microcomputer. All records are kept in paper files stored in boxes, and all "number crunching" is done by the external evaluator using his own hardware and software.

The April visit. The Federal Programs Coordinator was the only district person who participated in our site visit. The external evaluator was present only briefly the first day. A breakfast meeting was held with him the second day -- largely on his own (unpaid) time.

It was clear that evaluation resources in the district were extremely limited. The external evaluator's $2,500 contract required him to do all of the statistical data analysis plus six days worth of observing classrooms and providing feedback to project teachers. The district did not wish to alter this plan to allow time for reviewing the Users' Guide and participating in our meetings.

The Federal Programs Coordinator had read the Users' Guide prior to the March meeting. He claimed to agree with all of its recommendations although there were many he could not implement because of limited

resources and others that conflicted with Texas laws/regulations. He offered no suggestions for improving the guide.

He was particularly interested in the Users' Guide recommendations that curricular validity be a major factor in selecting tests to measure achievement growth. He claimed that the instruments used in Valley View were chosen because they were used in neighboring districts even though there were major differences between the districts' curricula. In thinking about the currently used tests, he felt that they might have low curricular validity. He indicated he would give high priority to reviewing alternative instruments for possible use next year.

We also discussed the problems associated with functional-level testing in Texas. Given the in-level testing requirement, it was mentioned that looking for tests with low floors as well as a good curriculum match might be a useful strategy.

No other significant issues were discussed during the March meeting.

The June Meeting. The June meeting was conducted after school had already closed. Only the Director of Federal Programs was available for discussion. Again, he had no specific comments on the Users' Guide.

We discussed the fact that Valley View's external evaluator would be implementing a TIERS Model A evaluation rather than a gap-reduction design. The RMC site visitor then explained that the two designs were closely related and that the former could be converted to the later with minimal computations on a hand calculator. The procedures were worked out and a step-by-step computational routine was developed. RMC also offered to provide technical assistance in making the conversion.

The discussion then turned to the "fairness" of using national norms as a comparison group for Valley View students. Pointing out that the comparison was neither fair nor unfair as long as no assumption was made

as to how well the project students would do without the project did little to convince the Federal Project Coordinator.

He would only feel comfortable using a local, mainstream comparison group -- and there is no such group in Valley View.

We discussed other possible comparison groups, but without reaching a satisfactory resolution.

Additional time was spent going over the curricular validity issue. This discussion appeared to be useful for reinforcing the conclusions that had been tentatively drawn during the previous visit, but no new ideas were introduced.

At the end of the first day, there appeared to be nothing left to discuss. The site visit was thus concluded one day ahead of schedule.

Valley Center, California

During the first visit to Valley Center, activities focused on analyzing the project and evaluation designs. The analysis revealed that project objectives needed clarification, and two concerns about the objectives were identified.

First, student outcome objectives used grade-equivalent scores as growth measures. Use of these scores, while appealing to practitioners and laymen, is psychometrically unsound. Project staff asked whether expected gains expressed in grade-equivalent scores can be converted to another metric such as NCEs. The Users' Guide should address this issue, as well as providing guidance on how to write student outcome objectives that reflect the gap-reduction design.

Second, curriculum-specific objectives are difficult to write. They tend to be confused with specific expectations, benchmarks, or skills,

such as "find the sum of two numbers." A curriculum-specific outcome objective should include a reference to certain instructional activities and/or materials which are (usually) designed to teach several skills. In addition, such an objective should specify the method of assessment (such as a unit test or checklist of skills) and the criterion for success. The Users' Guide needs to make this clear.

In analyzing the project design, it became apparent that different types of students having different degrees of English proficiency receive different types of instruction (see Appendix H). After an initial meeting with the project director, a meeting was held in which all project staff could discuss instructional design, what teachers should be doing in their classrooms, and other critical features of the project. While some initial thoughts were formulated, some confusion about the design of the project remained.

During his second visit to this site, Lam visited project classrooms and met again with project staff. Son after this meeting, the site developed a list of global and specific critical features. This list, included in Appendix G, closely follows Users' Guide recommendations. In generating the list, project staff expressed confusion about whether critical features were only instructional features or whether activities such as district support, parent involvement, or project management should be included. The Users' Guide should explain critical features more fully.

Based on the list of critical features, project staff have begun to develop a classroom observation instrument. This instrument, included in Appendix G, is not yet complete. As recommended by the BEES, project staff also generated evaluation questions and identified problems in program implementation.

In reviewing tests to determine their curricular validity, the project staff decided to review test objectives rather than individual items as recommended by the BEES. Project staff feel reviewing test objectives

is more efficient. Their experience suggests that the BEES might include a section on matching test and curricular objectives.

Project teachers reviewed the CTBS, identifying for each test objective whether or not they teach that objective at their grades. They also rate each objective in terms of the importance of developing the skills, the number of skills covered, and the amount of time needed to teach the skills (see Appendix H). We will examine these data to compare the adequacy of this objectives approach to the currently recommended item-by-item approach.

Data management is also a problem for Valley Center. Most data are maintained in separate paper files and have not been combined onto a central student form. Future efforts will be needed to create a computerized data management system. The Users' Guide should provide more guidance on computer software for data management.

San Jose Unified School District

Because of the proximity of the district to the RMC office, four, instead of two site visits were made to San Jose.

While there are two Title VII projects in the San Jose School District, only one of them, the Vietnamese and Other Indochinese Curriculum and Educational Services (VOICES), participated fully in the pilot testing.

During the first two visits to the district, the district evaluation and testing coordinator was present at the group meetings. The first concern the project staff had was the kind of data the project should report for evaluation (The Guide only describes what data to collect).

A list of data to report was subsequently generated and included in the later version of The Guide.

In the first two meetings, the project and evaluation designs were discussed. During that time, the project's needs assessment data, objectives, and instructional design were examined, clarified, and elaborated where necessary. Since the project was in its third year, there was some reluctance to make too many changes to the original designs. However, it was agreed that the project's actual instructional features should be identified and classroom observations should be conducted to monitor and determine the extent to which these desired instructional features are actually being implemented in the classrooms.

Through brainstorming among Lam and the project staff, some examples of specific critical instructional features were identified. Two of these features are: (a) aide summarizes, using L1, the content taught by teachers, and (b) students are grouped by reading ability. Using all of the examples, project staff were supposed to develop a checklist of specific critical instructional features for classroom observations. However, due to time and resource constraints, thus far the project has developed only a general observation form (see Appendix H). If the project receives an extension for two more years, a more specific observation form tailored for the project, which the project staff feel is important to have, will be needed to document program implementation.

A major component of the VOICES project is materials development. The project has been developing Vietnamese language arts and reading materials (programs) and social studies materials in both English and Vietnamese for kindergarten through fifth grade. For the VOICES project, the adequacy of the materials developed by the project is determined by examining student performance on the CTBS. What is missing is the intermediate step of reviewing the material for quality of construction and appropriateness. What the project staff found lacking in the Guide are procedures for evaluating materials development efforts. There is no guide on how to evaluate the quality of instructional materials either commercially or locally developed.

The VOICES project staff find curriculum-specific objectives easier to develop and more appropriate than do staffs of other projects since all VOICES teachers use the same curriculum, and there are criterion-referenced tests developed to accompany the curriculum. However, it is still not clear whether or not curriculum-specific objectives are needed if student skill areas and curricula are specified in advance and verified via classroom observation, and data on student classroom performance are collected. Writing curriculum-specific objectives may be an extra burden that can be avoided in this project.

Since the project receives adequate assistance from the district evaluation and testing office in data collection and management, not much technical assistance was given in this area.

During the last site visit, Lam visited a kindergarten class to gather first-hand information regarding classroom processes. Based on this observation, he was able to provide suggestions for improving the draft classroom observation form.

A problem encountered by the VOICES project that has not been dealt with in the Users' Guide is a concern about estimating student growth across time when the test is changed. For the VOICES project, there will be a change from a commercial test to a test tailor-made for the San Jose School District. How one should compute RGI's with different tests given at different points in time should be discussed in the Guide.

New York Community School District #5

There are two Title VII projects in New York Community School District #5. However, most of the pilot testing effort has been concentrated on the Critical Thinking Skills (CTS) project, which began in October of 1986. The CTS project is designed to enhance critical thinking skills of LEP students in grades K through 3 by instilling critical thirking teaching strategies into the bilingual classrooms. Teacher training and the

provision of teaching aides are the activities used to produce these results.

Lam worked with the project director, resource teacher, bilingual program coordinator, and external evaluator on ensuring the project's evaluability. The project objectives were completely revised in accordance with Users' Guide recommendations. The philosophy and theoretical framework underlying the teaching of critical thinking skills is very well defined. The challenge the project faces is to translate the theory into classroom practices, implement these practices, and then measure student growth in critical thinking skills.

Through a great deal of discussion, the project developed a classroom observation checklist that covers thinking skills teaching strategies and methodology for teachers and aides, and classroom management, environment, and interaction (see Appendix J). In the coming year, the checklist will be pilot-tested and then used to monitor and document the level of program instructional service delivery. The initial data collected can also be used to determine areas in which teachers or aides should receive training. This point is not mentioned in the current Users' Guide. As do instruments developed in other pilot sites, the observation instrument requires trained observers who are familiar with bilingual and ESL methodologies.

Another problem in teaching critical thinking skills is to assess student growth in these skills. The project staff has been searching for a satisfactory test, but has not found one. Consequently, the project is now developing its own test, which is against the recommendation of the BEES. Even if the quality of the test is acceptable, the project will have to administer the test to a comparison group in order to implement the gap-reduction analysis. Administering any test to any students in New York City for the sake of obtaining comparison data for a Title VII project may present some legal problems. This point should be acknowledged in the Users' Guide.

The project has been diligently collecting student data as recommended by the BEES, using the sample forms in the Volume Two of the Users' Guide. Currently, data have been collected from all Title VII students. The next step is to enter these data into the computer. Creating a computerized data base is an area in which the project needs technical assistance.

As in the other New York site, student mobility is a major problem for obtaining a representative sample of data at this site.

New York Community School District #10

Project CARE (Career Awareness Resource In Education) is a project in its first year. It is designed to infuse ESL methodology and career education into regular bilingual classes through teacher training and parent involvement and education. In an initial meeting with the project director and her supervisor, a great deal of time was spent discussing what the program consists of and how best to document its implementation.

The program focuses on career education for students and on the enhancement of teachers' ESL skills. Therefore, the evaluation staff needed to develop a checklist of critical features for these two areas. An ESL checklist was adapted from a district checklist, but it needs revision to make it easier to implement (see Appendix I). A checklist for career education has not yet been developed.

During the two visits to this project, project objectives were clarified and revised. There was also a discussion of how the project should measure changes in students' self concepts. Project staff were not happy with available paper and pencil measures. Lam suggested that project staff follow Users' Guide recommendations to use indirect indicators such as absenteeism or participation in extra-curricular activities. There is some question about whether one should expect the activities of this project to produce improved self concepts. The Users' Guide should have a

clearer discussion of the need for a logical linkage between treatment and anticipated outcomes.

The project also has a problem with high attrition. The posttest might include a completely different group of students from the pretest. The Users' Guide should address analysis issues related to high student mobility. The suggestion has been made that attrition at lower grade levels is more random than in the upper grades, so that systematic bias in analyses should be less at those grades.

New York state prohibits parents from observing classrooms, so the Users' Guide recommendation is inappropriate for this project. It is also not possible in New York for teachers to review tests item by item, since they are prohibited from seeing the tests except on the dates they are being administered.

III. SUMMARY OF CHAPTER-BY-CHAPTER REVIEW OF USERS' GUIDE

Comments from all Participants

Chapter I. Introduction

- The chapter should include a better description of the BEES, including its limitations.

- The chapter should make clear who the BEES is designed for and who is the audience for the Users' Guide.

- The chapter should include some discussion of previous research.

Chapter II. Ensuring the Project's Evaluability

- Needs assessment. The discussion is too technical. Give an example of a good needs assessment. Don't use a question format. If mobility is high, these data may not be appropriate baseline data. Need more discussion of baseline data.

- Objectives. Drop curriculum specific objectives. Define what objectives are needed to conform to what regulations. Realistic criteria are very difficult to determine.

- General. The chapter needs to be more definite. The format is too busy. The chapter needs a glossary, or definitions in the text. More examples and a clearer format are needed.

Chapter III. Planning the Evaluation

- It is difficult to form an evaluation team -- it is not always possible to involve other staff.

- It may not be possible to involve parents in an evaluation.

- This chapter should tell what implementation difficulties really are --use realistic examples.

- There are too many tables in this chapter -- they break up the text.

Chapter IV. Documenting Program Processes

- Tell exactly what is required by the regulations and what is an interpretation or recommendation by the authors of the Users' Guide.

- The Users' Guide should make clear that self reports should not be used without corroborating information.

- The examples in this chapter are poor. Use positive, not negative examples.

- Teachers' language fluency is very hard to assess.

- Other information would be more useful to collect than the information listed.

- Discuss who should conduct classroom observations.

- Many teachers will not agree to classroom observations.

- Differentiate between conducting classroom observations for teacher evaluation and for program evaluation.

- The Guide should make clear that a "delayed" report observation is only delayed until the observer has left the classroom.

- The data recommended for collection on students, instructional services and teachers should be prioritized.

As discussed previously, two questionnaires were administered, one prior to the June workshop, which addressed the availability and accessibility of these data. The second questionnaire was administered during the June workshop and it addressed the usefulness of these data for estimating program effectiveness[1]. Table 4 and Table 5 summarize results from these two questionnaires. As can be seen in Table 4, the information that has more "not available" responses than "available" responses include years of education missed, dates entered and exited other settings, enrollment in postsecondary education institution, and minutes per week devoted to specific learning activities.

In Table 5, it can be seen that information regarding whether or not the student is enrolled in public or private school and sex appear to have little importance for estimating program effectiveness. Other information that has questionable importance include date began U.S. residence and date of birth. All other data listed in the survey are perceived to be important or very important for estimating program effectiveness.

---

[1]Due to the similarity in response to the question regarding the usefulness of the data for estimating program effectiveness and the question regarding the usefulness of the data for improving program performance, only the responses to the first question were summarized

TABLE 4

Evaluation Assistance Center-West

Data Collection and Reporting Survey

This survey is designed to determine the extent to which the following list of information can be collected for Title VII local project evaluations.

Please rate each of the following list of variables in terms of availability and accessibility in your district.

	Information Available?				If Not Available, How Accessible?			
I. STUDENT INFORMATION	Y	N	DK	OTHER	MEAN	MIN	MAX	N
	(%)							
A. Background Information								
Date of birth	89	0	11	0	0	0	0	0
Place of birth	100	0	0	0	0	0	0	0
Ethnicity/language group	100	0	0	0	0	0	0	0
Date began U.S. residence (date)	44	22	22	11	2.5	2	3	2
Language used in home	100	0	0	0	0	0	0	0
Fluency in English as of (date)	100	0	0	0	0	0	0	0
Fluency in native language (L1) if other than English as of (date)	89	0	0	11	0	0	0	0
Years of education completed as of (date)	67	33	0	0	1.7	1	2	3
Years of education missed as of (date)	33	44	22	0	1.7	1	2	3
Family income status (e.g., AFDC, participation in National School Lunch Program)	67	33	0	0	3	1	5	3
Public or private school	78	11	0	0	4	4	4	1
School grade level	100	0	0	0	0	0	0	0
B. Service Data								
Dates entered and exited project	89	11	0	0	5	5	5	1
Other instructional settings (e.g., mainstream class, Chapter 1 program, special education, migrant program, ESL program)	89	11	0	0	3	3	3	1
Dates entered and exited other settings	22	67	11	0	2.5	1	4	6

	Information Available?				If Not Available, How Accessible?			
	Y	N	DK	OTHER	MEAN	MIN	MAX	N
	(%)							
C. Non-Test Data								
Years retained in grade as of (date)	89	11	0	11	1	1	1	1
Whether or not student dropped out	67	33	0	0	2.3	2	3	3
Days absent from project classroom as of (date) and total school days possible as of that date	100	0	0	0	0	0	0	0
Date of special education referral or placement	89	11	0	0	2	2	2	1
Date of placement in gifted and talented program	67	22	0	11	2	2	2	1
Enrollment in postsecondary education institution	11	56	11	22	1.6	1	3	5
Setting from which entered project	67	22	11	0	1	1	1	2
Grade point average prior to entry in project	33	33	11	22	3.7	2	5	3
Setting to which exited from project	44	44	11	0	2	1	4	4
Grade point average after exit	22	22	0	11	3.5	2	5	4
D. Former Project Students								
Standardized achievement test data	78	22	0	0	3	3	3	2
Classroom performances (e.g., grade)	89	11	0	0	5	5	5	1
II. INSTRUCTION								
Academic and non-academic subjects taught in the project	100	0	0	0	0	0	0	0
Method of service provision (e.g., tutoring, classroom instruction)	89	0	0	11	0	0	0	0
The following variables apply to each class or								
English and L1 materials and how they are used	78	22	0	0	3	3	3	2
Percents of instruction conducted in L1 and in English	67	33	0	0	2.3	1	3	3

	Information Available?				If Not Available, How Accessible?			
	Y	N	DK	OTHER	MEAN	MIN	MAX	N
Pattern of classroom language use (e.g., concurrent, preview/review, or alternating)	67	33	0	0	3	2	4	3
Critical teaching strategies including instruction and management procedures (e.g., use of active teaching behaviors, use of content from students' home culture)	56	44	0	0	3	2	4	4
Locations of instruction (e.g., mainstream classroom, special bilingual classroom, newcomer center, resource room)	100	0	0	0	0	0	0	0
Student groupings (e.g., LEP students are grouped by English proficiency, all LEP students are together all day)	78	22	0	0	2.5	2	3	2
Student/teacher ratio, student/aide ratio per class	100	0	0	0	0	0	0	0
Variability of students within a classroom (e.g., age, grade, educational attainment, home language, language abilities)	89	11	0	0	3	3	3	1
Function of aides in classroom (% instruction, % management)	67	33	0	0	3.7	2	5	3
Form of instructional group (e.g., whole class, tutorial)	89	11	0	0	4	4	4	1
For each subject area taught, minutes per week devoted to specific activities (e.g., math workbooks, English spelling, language lab)	44	56	0	0	1.7	1	3	4
III. STAFF								
Academic preparation of each staff member, including:								
• level of education,	78	22	0	0	3.5	2	5	2
• academic preparation in field of education,	67	22	11	0	3.5	2	5	2
• credentials and certificates, and	89	11	0	0	2	2	2	1
• academic preparation in field of bilingual/bicultural education	67	22	11	0	3.5	2	5	2

	Information Available?				If Not Available, How Accessible?			
	Y	N	DK	OTHER	MEAN	MIN	MAX	N
Teaching experience of each staff member, including:								
• years of experience in monoligual classrooms, and	67	33	0	0	3.3	2	5	3
• in bilingual classrooms	67	33	0	0	3.3	2	5	3
Language abilities of each staff member, including languages understood and spoken and degree of fluency in:								
• English	56	44	0	0	2.7	2	3	4
• L1	56	44	0	0	2	4	3	4

TABLE 5

STUDENT, INSTRUCTION, AND STAFF INFORMATION
(Response to question: "How important is information for estimating program effectiveness?")

I. STUDENT INFORMATION

A. Background Information	Mean[1]	SD[2]	N	Min	Max	1	2	3	4	5
						(Percent)				
Date of birth	3.4	1.1	9	2	5	22	22	22	22	22
Place of birth	3.6	1.4	9	1	5	11	11	22	22	33
Ethnicity/language group	4.2	1.1	9	2	5	0	11	11	22	56
Date began U.S. residence	3.1	1.8	9	1	5	11	22	33	11	22
Language used in home	4.8	0.7	9	3	5	0	0	11	0	89
Fluency in English as of	4.8	0.7	9	3	5	0	0	11	0	89
Fluency in native language (L1) if other than English as of	4.8	0.7	9	3	5	0	0	11	0	89
Years of education completed as of	4.4	0.7	9	3	5	0	0	11	33	56
Years of education missed as of	4.6	0.7	9	3	5	0	0	11	22	67
Family income status (e.g., AFDC, participation in National School Lunch Program)	3.7	0.9	9	3	5	0	0	56	22	22
Public or private school	2.6	0.9	8	1	4	13	25	50	13	0
School grade [illegible]l	4.4	0.9	9	3	5	0	0	22	11	67
Sex	2.6	1.4	9	1	5	33	11	33	11	11
B. Service Data										
Dates entered and exited project	4.9	0.3	9	4	5	0	0	0	11	89
Other instructional settings (e.g., mainstream class, Chapter 1 program, special education, migrant program, ESL program)	4.7	0.7	9	3	5	0	0	11	11	78
Dates entered and exited other settings	4.1	0.9	9	3	5	0	0	33	22	44

C. Non-Test Data	Mean[1]	SD[2]	N	Min	Max	1	2	3	4	5
						(Percent)				
Years retained in grades as of	4.3	.9	9	3	5	.0	0	22	22	56
Whether or not dropped out	4.7	.7	9	3	5	0	0	11	11	78
Days absent from project classroom as of -- and total school days possible as of that date	4.4	.7	9	3	5	0	0	11	33	56
Date of special education referral or placement	4.3	1.0	9	2	5	0	11	0	33	56
Date of placement in gifted and talented program	3.9	1.3	8	2	5	0	25	0	38	38
Enrollment in postsecondary education institution	3.7	1.7	9	1	5	22	0	11	22	44
Setting from which entered project	4.0	1.5	9	1	5	11	11	0	22	56
Grade point average prior to ntry in project	3.6	1.2	9	1	5	11	0	33	33	22
Setting to which exited from project	3.8	1.5	9	1	5	11	11	11	22	44
Grade point average after exit	4.2	.8	9	3	5	0	[illegible]	22	33	44
D. Former Project Students										
Standardized achievement test data	4.9	.3	9	4	5	0	0	0	11	89
Classroom performance (e.g., grade)	4.7	.5	9	4	5	0	0	0	33	67
II. INSTRUCTION										
Academic and non-academic subjects taught in the project	4.9	.3	9	4	5	0	0	0	11	89
Method of service provision (e.g., tutoring, classroom instruction)	4.9	.3	9	4	5	0	0	0	11	89

The following variables apply to each class or subject matter offered by project:

	Mean[1]	SD[2]	N	Min	Max	1	2	3	4	5
						(Percent)				
English and L1 materials and how they are used	4.7	1.0	9	2	5	0	11	0	0	89
Percents of instruction conducted in L1 and in English	5.0	0	9	5	5	0	0	0	0	100
Pattern of classroom language use (e.g., concurrent, preview/ review, or alternating)	4.6	.7	9	3	5	0	0	11	22	67
Critical teaching strategies including instruction and management procedures (e.g., use of active teaching behaviors, use of content from students' home culture)	4.8	.4	9	4	5	0	0	0	22	78
Locations of instruction (e.g., mainstream classroom, special bilingual classroom, newcomer center, resource room)	4.2	.9	9	2	5	0	11	0	44	44
Student groupings (e.g., LEP students are grouped by English proficiency, all LEP students are together all day)	4.7	.5	9	4	5	0	0	0	33	67
Student/teacher ratio, student/ aide ratio per class	4.6	.7	9	3	5	0	0	11	22	67
Variability of students within a classroom (e.g., age, grade, educational attainment, home language, language abilities)	4.7	.7	9	3	5	0	0	11	11	78
Function of aides in classroom (% instruction, % management)	4.7	.7	9	3	5	0	0	11	11	78
Form of instructional group (e.g., whole class, tutorial)	4.9	.3	9	4	5	0	0	0	11	89
For each <u>subject area</u> taught, minutes per week devoted to specific activities (e.g., math workbooks, English spelling, language lab)	3.6	1.2	8	1	5	22	11	11	0	44

III. STAFF

	Mean[1]	SD[2]	N	Min	Max	1	2	3	4	5
						(Percent)				
Academic preparation of each staff member, including:										
• level of education,	4.4	.9	9	3	5	0	0	22	11	67
• academic preparation in field of education,	4.8	.7	9	3	5	0	0	11	0	89
• credentials and certificates, and	4.7	.7	9	3	5	0	0	11	11	78
• academic preparation in field of bilingual/bicultural education	5.0	0	9	5	5	0	0	0	0	100
Teaching experience of each staff member, including:										
• years of experience in monolingual classrooms, and	4.3	1.0	9	3	5	0	0	33	0	67
• in bilingual classrooms	4.8	.4	9	4	5	0	0	0	22	78
Language abilities of each staff member, including languages understood and spoken and degree of fluency in:										
• English	5.0	0	9	5	5	0	0	0	0	100
L1	5.0	0	9	5	5	0	0	0	0	100

[1] 1=least important; 5=most important

[2] SD=Standard deviation

Table 6 summarizes results from the questionnaire designed to obtain input from the field site staff in terms of accessibility and usefulness of the project background information recommended by the BEES. As can be seen in the table, the only two community characteristics that are perceived as accessible and useful are percentage of residents below federal poverty level and ethnic makeup of community. All the data regarding district demographics and school characteristics are considered accessible and useful.

Chapter V. Selecting/Adapting/Developing Instruments for Assessing Summative Outcome Objectives

- It is very difficult to find curriculum-relevant standardized tests. It is also hard to get a group to review tests item-by-item.
- Why not write a test in L1 or translate a test to L1? If there are good reasons, they should be spelled out.
- Why use reliability as a test selection criterion when all standardized tests have similar reliability?
- Why use cultural bias as a test selection criterion when it is not important for estimating growth?
- Give examples of cultural bias. (An example was suggested.)
- Telling new teachers to teach test-t king skills can cause problems.

Chapter VI. Implementing the Gap-Reduction Design

- This chapter is too easy for evaluators and too hard for practitioners.

TABLE 6

Project Background Information

	Accessibility					Usefulness				
	Mean[1]	SD	N	Min	Max	Mean[2]	SD	N	Min	Max
A. COMMUNITY CHARACTERISTICS										
Percentage of residents below federal poverty level	3.4	1.6	9	1	5	4.2	1.3	9	1	5
Unemployment rate	2.7	1.6	9	1	5	2.8	1.6	9	1	5
Workforce composition	2.2	1.0	9	1	3	2.9	1.4	9	1	5
Ethnic makeup of community	4.1	1.2	9	2	5	4.0	1.3	9	1	5
Average length of residence in community	2.0	1.1	9	1	4	3.1	1.5	9	1	5
Percentage of single-parent families	1.9	.9	9	1	3	3.1	1.3	8	1	5
Educational level of adult community residents	1.3	.7	9	1	3	2.8	1.7	9	1	5
Community attitude toward bilingual education	2.1	1.1	9	1	4	3.9	.8	9	3	5
B. DISTRICT DEMOGRAPHICS										
Approximate peak student enrollment	4.8	.4	9	4	5	4.2	1.0	9	2	5
Number of schools in district	4.9	.3	9	4	5	4.2	1.0	9	3	5
Characterization of school district (rural, inner city)	4.9	.3	9	4	5	4.6	.7	9	3	5
Number of classrooms in bilingual program	4.9	.3	9	4	5	4.9	.3	9	4	5
Percentage of students eligible for free lunch	4.7	.5	9	4	5	4.4	1.0	9	2	5
Ethnic makeup of student body	4.6	.7	9	3	5	4.7	.7	9	3	5
C. SCHOOL CHARACTERISTICS										
Student enrollment	4.9	.3	9	4	5	4.7	.5	9	4	5
Percentage of students eligible for free lunch	4.8	.4	9	4	5	4.6	1.0	9	2	5
Ethnic makeup of student body	4.6	.7	9	3	5	4.7	.7	9	3	5
Number of bilingual teachers	4.9	.3	9	4	5	5.0	0	9	5	5
Number of bilingual aides	4.8	.7	9	3	5	5.0	0	9	5	5
Number of bilingual classes	4.9	.3	9	4	5	5.0	0	9	5	5

1 - 1 = not accessible; 5 = very accessible
2 - 2 = not useful; 5 = very useful

- There should be more discussion of locating appropriate comparison groups, especially for classes such as kindergarten or special ed.

- Regression should be explained better. Avoiding regression bias requires too much testing. What is spuriously inflated? How much?

- It is not possible to administer early posttests to students when you do not know they are leaving.

- How long should a student participate in the project over the year to be considered a participant?

- Discussion on missing data is misleading.

- Explain the differences and similarities between the Gap-Reduction Design and the Norm-Referenced Design.

Chapter VII, Collecting Outcome Data

- The expectations in this chapter are unrealistic for a large, urban project.

Evaluator Comments on Chapters VI and VIII

The two chapters of the Users' Guide that deal with implementing the gap-reduction design (Chapter VI) and processing and analyzing data (Chapter VIII) were discussed separately with conference attendees who considered themselves to be evaluators and those who considered themselves to be project directors/administrators. What follows are the comments and suggestions of the evaluators.

Chapter VI. Implementing the Gap-Reduction Design

None of the evaluators had any problems with the content of the chapter, although there was considerable discussion of the RGI metric. Earlier in the conference it had been noted that RGIs become meaningless if the comparison group makes zero growth or negative growth. While this situation should never arise with achievement test scores (unless they are expressed in NCEs -- a conversion that reduces norm-group growth to exactly zero), it could arise with non-test data such as attendance and with other measures including classroom grades where growth is not expected.

A discussion of these issues is not currently in the Users' Guide, and the consensus was that it should be. It was also suggested that the purpose of the RGI metric (to provide a comparable "score" useful for making comparisons between different projects using different tests and different comparison groups) be explained. It should also be explained that where a metric does not have to perform equating functions, evaluators need not compute RGIs or even standardize growth estimates. Under these conditions, the amount of gap-reduction itself provides all of the needed information.

Keith Baker suggested a refinement of the gap-reduction measure for possible inclusion in the BEES. This refinement would express the amount of gap reduction as a percentage of the pretest gap. It would be particularly informative if the mainstream classmates of the project participants were used as the comparison group. A 100% gap reduction would then mean that no gap remained and that the intervention's job was complete.

Following the technical discussions just described, we worked through some hypothetical examples of how the gap-reduction design could be used with non-test data.

Chapter VIII. Processing and Analyzing Data

The evaluators reviewed chapter VIII near the end of the conference under considerable time pressure. We did not get past p. VIII-6.

The first issue addressed was score conversions (p. VIII-2). The Users' Guide presents this as a necessary step in the data processing procedure. Elsewhere, however, it says that raw scores may be used. The Users' Guide should make clear when conversions are required and when they are not. Why are raw scores less desirable? In general, a discussion in the Users' Guide is needed to explain the different types of scores and how they should be used for analysis and reporting.

The construction and management of data bases was another concern. It was felt that not enough guidance is presented on this topic. It would be nice if the Users' Guide could present information about different computer software--what will work on what computer, what each package can and cannot do, and how much each costs. It would also be nice to know how difficult/easy each package is to use. Most people will need more specific guidance than is presented. A flow diagram would also help.

It was suggested that we give more emphasis to recording student ID numbers and grade levels on all data collection instruments and files.

One evaluator felt that classroom grades were too unreliable to use for monitoring the progress of reclassified LEP's. He described classroom grades as "just one option and a bad one at that." He felt we should recommend the use of professionally-developed tests whenever possible. He also felt that ratings or rankings made by classroom teachers would be as unreliable as classroom grades.

We were just beginning to discuss Other Summative Outcome Analyses when time was called. It was clear, however, that the Users' Guide discussion of baseline data was confusing to some and needs work.

IV. LESSONS LEARNED FROM THE FIELD TEST

The lessons learned from the field can be roughly categorized into two groups, those that pertain to the system itself and those that pertain to the Users' Guide. Each of these categories is addressed below.

Lessons Related to the BEES

In general, the sites' reactions to the BEES were positive. Both during the course of our site visits and during the first day of the Exit Conference, there were many comments relating to the evaluability assurance component of the system. Between the Users' Guide and the technical assistance provided during site visits, most of the field test participants felt that their planning and evaluation activities had been significantly improved. BEES recommendations helped them to state their project objectives more clearly, to design project activities so that they related more directly to the objectives, and to focus their evaluation questions more sharply.

All sites found the recommended documentation and supplementary analyses excessively burdensome. They would like more direct guidance as to which of our recommendations are "driven" by the regulations and which are not. Among those not tied directly to the regulations, they would like to see some prioritization.

Most of the sites had one or more problems with our recommendations regarding test selection. Some have no choice in the matter and have to live with the state- or district-mandated instrument. Others are similarly precluded from functional-level testing, from modifying test instructions or time limits, or from making any changes whatsoever to the test items. Even when our recommendations could be implemented, some (e.g., reviewing candidate instruments item-by-item) were considered too consuming of time and effort.

The sites were also very positive about the gap-reduction approach to assessing project impact. Several evaluators expressed the opinion that the design was the only one that made sense and was feasible to implement. On the negative side, several districts found the required computations frighteningly complex. They would like us to provide user-friendly software that would free them of all the computational burdens (including interpolation/extrapolation and correcting for regression).

Data base management is an area where all but one site felt that the BEES' recommendations are inadequate. If nothing more, we should describe existing software that they might be able to use.

Project directors and teachers believe that classroom observation is very important to use in program evaluation. It seems likely that these staff members will conduct most, if not all of the observations.

Except for information regarding community characteristics, most of the data the BEES recommends collecting for evaluation and reporting appear to be accessible and potentially useful to project staff.

The BEES does not have enough emphasis on evaluating parent involvement; staff training, and materials development.

To supplement the preceding, somewhat subjective impressions, seven general questions regarding the BEES were posed to each of the nine field-test sites on the final day of the Exit Conference. These questions and the written responses of site personnel are presented below

Question 1. To what extent has the system aided the local projects in defining and objectively measuring the impact of the bilingual program on its students?

Four sites responded that the BEES had been helpful by providing a system in which the relationship between needs, objectives, data collection, and

monitoring of effectiveness is clearly defined. Two sites said that it was too early to tell at their sites, two sites were vaguely positive, and one site said that the gap-reduction model should be useful.

Question 2. To what extent can the system serve to improve local evaluation practices?

Two districts responded that the BEES had helped them make use of district evaluation data and other existing data in their evaluations. Three districts gave general but positive responses, and two said the system would helpful only to the extent that funds are available to implement it. One district said the BEES could improve all local evaluations involving minority students, and one said the system had been helpful in encouraging the project director to participate in the evaluation.

Question 3. To what extent can the system generate data suitable for use in local decision-making?

Two districts said that the information would be useful in gaining school board understanding, while a third district said the information would be useful at the project level but not at the district level. One district said the data generated would depend on the availability of resources to implement the system, and three districts gave general but positive responses.

Question 4. How adequately can the evaluation designs be integrated with a district's regular instructional and testing programs?

Four districts said the evaluation design might or would be very difficult to integrate with district practices. Three districts said it would fit in quite well, and two said it was too early to say.

Question 5. How suitable are the data generated by the system for synthesis across Title VII programs?

Three districts said outcomes should not be combined or compared across different types of projects and populations. Three said results would be suitable, and two said they would be suitable only if the RGI is actually workable. One district said they had been hoping for much more emphasis on data integration and synthesis, including software.

Question 6. What are the costs incurred in implementing the evaluation designs described in the Users' Guide?

- Costs associated specifically with the project site (e.g., specific to setting up and maintaining the records that the system requires);
- Start-up costs (e.g., training data collectors); and
- Ongoing costs of evaluation (e.g., collecting and analyzing data that are collected).

Two districts estimated costs in dollars. One estimated site costs of $30,000, start-up costs of $5-7,000, and ongoing costs of $6-7,000. A second district estimated $1,000 site costs, $1,000 start-up costs, and $8,000 ongoing costs. Five districts said that costs would be high, with three adding that at least one full-time person would be required. Two districts did not know.

Question 7. To what extent would a district require technical assistance in implementing the system?

Two districts said a full-time evaluator would be more practical than technical assistance. Six districts indicated that they would need help in data management and analysis, possibly through the provision of appropriate software. One district said a little technical assistance would be necessary but did not specify the area of the assistance.

Lessons Related to the Users' Guide

The Users' Guide has already undergone several major revisions. Clearly, however, substantial room for improvement remains. Perhaps its most significant shortcoming is its failure to communicate effectively with the disparate audiences. Its more technical parts are viewed as incomprehensible and full of jargon by program people. Evaluators, on the other hand, find the non-technical parts irrelevant to their task and the technical parts excessively discursive. Both technical and non-technical readers perceive the writing as a compromise that fails to serve either group as well as it might.

Throughout the Guide, there is a consensus that the writing style and the format could be improved. There are inconsistencies of style, headings, structure, and punctuation that need to be fixed. With somewhat less consistency, field test personnel and other reviewers made the following suggestions:

- the Users' Guide should include a glossary of terms.
- recommendations should be stated more directly
- more rules of thumb should be given
- more step-by-step procedures should be given
- there should be more examples
- there should be more explanation
- the document should be shorter
- there should be less explanation

- there should be separate Guides for project directors and evaluators

We have even received such radical recommendations as, "limit the Users' Guide to the gap-reduction design and discard everything else." One reviewer suggested just the opposite: "throw out all the outcome evaluation material and limit the Guide to process evaluation.

Conclusions

It will obviously be impossible to incorporate all of the recommendations into a revised BEES and Users' Guide as some of these are mutually contradictory. Many changes will be made, however, to make the system and the Users' Guide better meet the needs of their audiences. In approximate decreasing order of priority, we plan to do the following:

- provide software that will perform all of the gap-reduction computations
- develop an abbreviated Users' Guide that will present recommendations succinctly and without any unnecessary explanations
- present parallel conceptual (for project director) and technical (for evaluators) segments of the Users' Guide whenever psychometric and statistical issues must be addressed.
- develop some recommendations for data-base-management software
- separately identify required, important but not required, and probably useful categories of data to be collected and analyses to be done
- include a glossary of terms (already nearly completed)

- make our recommendations more directive and procedural
- include more examples and rules of thumb
- undertake a major editing and reformatting revision of the Users' Guide using professional document-design specialists.

It is not yet clear how we will respond to comments on the test selection/modification issue. Clearly we need to alter what we have, but we need to avoid conveying the impression that any old test will do. Perhaps we have been overly zealous in our recommendations regarding functional-level testing and modifying existing instruments. On the other hand, users of the BEES must be made aware of the fact that test floor effects and low curricular reliability can only make their projects appear less effective than they really are and that other measures of effectiveness must be sought if state or district mandates force either of these ill-fated circumstances upon them.

APPENDIX A

Site Selection Criteria

There is one, overriding criterion for the selection of field test sites of the "Bilingual Models" study. That criterion relates to the need to try out all parts of the User's Guide and associated materials.

Most parts of the Users' Guide will be applicable to all bilingual eduction project evaluations and thus will have no particular implications for field-test site selection. One very important exception to this general rule, however, is that portion of the Guide that deals with dividing observed growth into treatment-related and non-treatment-related components. Three alternative designs will be offered as means of accomplishing that objective (quasi time series, regression-discontinuity, and non-equivalent comparison group). We anticipate that many projects will find it impossible to implement any of these designs. Since they will be part of the Users' Guide, however, it will be our highest site-selection priority to find at least one field-test site that can and would be wiling to tryout our procedures for implementing each of these designs. It would be desirable to find two or three sites that would implement each design.

Our second-priority site-selection criterion will be diversity in the extent of projects' instructional usage of L1. Because different types of tests and testing procedures will be required, we would like to include projects that encompass L1 literacy as an academic subject, projects that use L1 as the medium of instruction for at least part of the curriculum (but do not teach L1 as a subject), and projects that use English as the medium for all instruction.

Our third-priority criterion relates to the linguistic characteristics of the students served, as different characteristics impose different requirements on both instructional and evaluative processes. We would like to include projects serving Hispanic students, projects serving Asian students, and projects serving students with heterogeneous language backgrounds. Our preference would be to include at least one project serving Hispanic students of Mexican origin, at least one project serving Puerto Rican students, and at least one project serving Cuban students. Among

projects serving Asian students, we would like to include at least one serving refugee students and one serving non-refugees as we again expect that the two groups would pose different problems for both educators and evaluators.

We would also like our sample to include projects of different sizes, projects serving different grade levels, projects whose evaluators have different levels of technical sophistication, and projects from different parts of the country. With only nine districts in our sample, of course, we may not be able to achieve all of these objectives. It will help, however, if we can find districts that have more than one project.

Finding districts for the sample and getting them to participate will probably both be non-trivial tasks. Our tentative plan is to begin by contacting the bilingual program directors in states with large numbers of programs. Since California, New York, and Texas have something on the order of 85% of all the programs in the country, we propose to begin with them. After explaining our objectives for the field-test sample, we will seek nominations. While following up on the nominations received from these states we will also seek nominations from other states, from the Evaluation Assistance Centers, and possibly from OBEMLA.

We will attempt to obtain a large number of nominations, both to allow for sites that will elect not to participate and to provide us with flexibility in filling our sampling frame. Before we begin contacting sites we will attempt to determine the characteristics each of them possess vis-a-vis our selection criteria. We will then begin soliciting their participation starting with the sites that have the most desirable characteristics.

The characteristics of the first site that agrees to participate will cause us to adjust the order in which subsequent sites are contacted. Similar reorderings will occur as each successive site is added to the sample. This approach, although cumbersome, is the one we believe will produce the best possible sample.

APPENDIX B

Protocol Invitation for Field Test Participation

Dear Superintendent:

This letter is a formal invitation to the _____ school district to participate, as one of nine school districts across the country, in a field test of the bilingual education evaluation system we have been developing under a contract with the U.S. Department of Education (see attached project summary). As I have discussed briefly with you, and more extensively with members of your staff, the new system is being developed in response to the widely recognized need to improve and standardize procedures for evaluating bilingual education interventions.

The ultimate goal of the system is to enhance the quality and effectiveness of the services being provided to limited-English-proficient students. For this reason, it has a heavy emphasis on process evaluation and on the integration of process and outcome information for the purpose of generating recommendations aimed at improving local projects. The purposes of the field test are to assess the feasibility of implementing the system as we have developed it and to identify changes that could be made to enhance its ease of implementation and to increase the utility of its information return.

Agreeing to participate in the field test implies that your staff will make a conscientious effort to implement our recommended evaluation practices to whatever extent situational and budgetary constraints will allow. We will not be <u>conducting</u> the evaluation in your district, only assisting responsible persons in the district to try out the procedures and materials we have developed.

Participation in the field test will represent some burden for your district. We believe, however, that this burden will be more than offset by the benefits that will accrue to your district through participating. The evaluation expertise of your staff will be substantially enhanced by virtue of the training and technical assistance we will provide. Furthermore, by the end of the year, you will have established sound evaluation

practices that will stand you in good stead over future years. During the field-test year itself we will also be able to assist with data analysis and interpretation and perhaps provide you with computer software that will also be useful in the future. The greatest benefit of all, of course, will be program improvement if the new system does indeed prove useful for that purpose.

As previously planned, your involvement in our filed test will involve sending your Title VII project director and one or two other persons associated with bilingual program evaluation to our office for one week of training. We will pay their travel and related expenses during that week, but cannot pay salaries or honoraria. The training will occur not long before, or shortly after the school year begins, depending on the availability of the persons to be trained.

During the school year, members of our staff will visit your project probably six times before and/or during critical evaluation activities. At those times we will monitor evaluation activities and provide technical assistance. Our purpose will not be to evaluate the manner in which your evaluation activities are being conducted but to identify deficiencies in the procedures and materials we have developed so that they can be improved. At the end of the year we will conduct a debriefing.

During the course of all our visits we will need to have access to your project director and evaluator(s) as we will need to discuss the adequacy and usefulness of our materials with them. We will attempt to be unobtrusive, however, and to conduct our discussions on a non-interference basis. We will, of course, treat all data provided by your district as confidential and will take whatever steps are necessary to protect the source of all information we collect.

As we are eager to begin by sending materials to your staff for review, I hope that you will let me know your decision as soon as possible. Should you have any questions, I will, of course, be pleased to try to answer them.

Sincerely yours,

G. Kasten Tallmadge
Senior Vice President

GKT:mb
Enclosure

APPENDIX C

Agendas and Correspondence Relating to the Two Project Workshops

Tony

SRA Technologies, Inc. 2570 West El Camino Real Suite 610 Mountain View, CA 94040 415-941-7084

27 October 1986

Norma Anzaldua
Valley View Independent School District
Rt. 1, Box 122
Pharr, TX 78577

Dear Norma:

During the last day of our workshop, I believe I promised to provide you with specific guidance in field testing the Users' Guide. After my first phone call to you and the others, I felt I had better keep my promise after all.

In general, what I have in mind for you to do is to have you and your evaluator (and other involved individuals) read over the Users' Guide to determine what recommendations can be implemented and what recommendations cannot, and why. The next step is to make plans to carry out these "do-able" recommendations. After that, implement the identified recommendations and keep records of progress. Our role is to act as facilitators and to document all these activities.

Let me highlight some of the key things that you may want to consider doing at this point in time. They are:

1. Review your proposal and evaluation report with your evaluator and relevant staff for the purpose of clarifying and modifying (if needed) your needs statements, project objectives, and project plan following the guidelines presented in the Evaluation Focusing chapter. Come up with some evaluation questions if you don't have any, or modify your existing evaluation questions if they do not meet the criteria listed in the Focusing chapter.

2. Together with the relevant "others," determine your project's global and specific critical program features which you can observe and/or document via other procedures such as questionnaires or interviews. Plan all these activities now.

3. For the new projects, plan your evaluation following the suggestions presented in the Evaluation Planning chapter. For continuation projects, try to incorporate some of the suggestions for this year's evaluation that are not too late to do. For example, involve more people in the evaluation process, construct an evaluation activities chart (see the two examples I gave in the workshop), anticipate potential obstacles and pitfalls, and implement some of the strategies for increasing the utilization of evaluation findings.

4. For the new projects, perhaps you may have the luxury of selecting, adapting, or developing your assessment instruments. If that's true, use Chapter 4's recommendations as your guidelines. For the continuation projects, it will be useful to review the quality of the instruments that you are stuck with using the test selection criteria described in Chapter 4. You may want to consider modifying your test as suggested in Chapter 4. You may want to develop instruments for assessing program implementation. Read Chapter 4. Is it useful? Are the recommendations do-able? In Chapter 4, the recommendations not to use questionnaires and interviews should be interpreted as using them with care.

5. Remember the workshop Day One exercise (attached)? That awful list of variables that you were asked to look over and rate. Make plans to collect them now. I'd like to know why you can't collect some of these data. I'll be sending you actual forms that you may want to use.

6. Start a routine system for recording relevant events and circumstances that relates to problems and difficulties in implementing program and evaluation activities (see Chapter 3, III-5 to III-6). You, your teachers, and (perhaps) the evaluator should keep a log of this type of information, which should be consolidated periodically (e.g., every staff meeting).

7. Some of you may have to begin making plans for constructing your data management system. Consult Chapter 5 (V-22).

8. Make plans to collect data following recommended procedures presented in Chapter 5 (V-1, V-15).

The above are some of the things that come across my mind. There is nothing better than reading the Guide carefully and coming up with your own list. The purpose of my first visit will be to help you to finalize your agenda of things to do for the field test, things you can do to improve your evaluation and your program. From time to time, I'll be sending you parts of the Guide that have been updated or revised or added.

As I mentioned to you over the phone, most likely we will have another "get together" sometime in late June or early July. At that time you and the other sites will present what you did and did not do and why. After individual presentations, we will then discuss as a group or subgroups the merits of the evaluation system (BEES) following the seven questions listed in the workshop handout labeled, "Purpose of the Field Test" (attached) and the content of the Users' Guide following the six points listed in another handout labeled "Open Discussion: Content of the Users' Guide" (attached). Throughout the year when you're working with the Guide, it will be useful to keep in mind the above issues that are the foci of the field test.

As you may have gathered by now, we're really asking a lot! Yes, but that's because we really value your input which we believe can significantly improve the quality of the Guide. After all, you are the "chosen" ones! Your district, your project, and most importantly, your name will be remembered forever in the annals of bilingual education evaluation! Just think, this clout, plus your enhanced skills in proposal writing and evaluation will guarantee endless awards on your future proposals.

Talking about proposals, I think it will be a great idea if you would use the Guide to help you write your next proposal. Not only will that create an opportunity for you to study the Guide, I also believe that the quality of your proposal should be enhanced (that means a better chance in winning the big bucks!)

Jokes aside, we really appreciate your help. I believe if we all give it our best shot, the field testing will be a rewarding experience for you and us. I'll make sure the benefits are mutual and the work is as much fun as possible.

Talk to you soon.

Sincerely,

Tony C.M. Lam
Project Director

TL:TH
Enclosure

cc: Ray Morales

Sent to the 9 sites

13 April 1987

Cheryl Crawley
Salem-Keizer Public Schools
1309 Ferry Street SE
P.O. Box 12024
Salem, OR 97309

Dear Cheryl:

Time flies when you are having fun. Before you know it, the June meeting will be around the corner and the pilot testing of the Bilingual Education Evaluation System (BEES) will be history. I feel it is appropriate at this time to suggest to you things you may want to wrap up from now until June in preparation for your presentation in the upcoming meeting in Washington, DC.

To reiterate what we've discussed in the past, your role is to provide us with feedback from your group's perspective as practitioners regarding strengths and weaknesses of the BEES' Users' Guide. You should review the February 20, 1987 version of the Users' Guide carefully and pilot test its recommendations in two ways. The first (preferred) way is actual implementation of the recommendations, e.g., rewriting the project objectives. The second way is to review the recommendations that you are unable to implement. Then you should tell us which recommendations appear sound and implementable and which do not and why.

There are four things I'd like you to try out. They are: (a) revise project objectives, (b) develop evaluation questions, (c) collect process and outcome data suggested by the Guide (including descriptive information regarding problems associated with program implementation and evaluation activities), and (d) review existing test instruments (at least for one grade for one subtest, e.g., Reading Comprehension).

I've enclosed a copy of the data collection survey with this letter. If you have not completed it, please do so at your earliest convenience.

If you have any questions, you know where to find me and Kast. If I'm not in the office, leave a message and I'll get back to you.

Our project has hired a consultant who can assist you in regard to data management. Her name is Camille Harder and she can be reached at the same phone number as mine. She'll be in touch with you soon.

I know it's a lot of work but I believe the rewards should be just as plenty! Take care and I look forward to seeing you again.

Sincerely,

Tony Lam
Project Director

TL:TH
Enclosure

Evaluation Assistance Center-West

Data Collection and Reporting Survey

This survey is designed to determine the extent to which the following list of information can be collected for Title VII local project evaluations.

Please rate each of the following list of variables in terms of availability and accessibility in your district.

	Information Available?			If Not Available, How Accessible?				
I. STUDENT INFORMATION	Yes	No	Don't Know	Very Difficult				Very Easy
A. Background Information								
Date of birth	Y	N	DK	1	2	3	4	5
Place of birth	Y	N	DK	1	2	3	4	5
Ethnicity/language group	Y	N	DK	1	2	3	4	5
Date began U.S. residence (date)	Y	N	DK	1	2	3	4	5
Language used in home	Y	N	DK	1	2	3	4	5
Fluency in English as of (date)	Y	N	DK	1	2	3	4	5
Fluency in native language (L1) if other than English as of (date)	Y	N	DK	1	2	3	4	5
Years of education completed as of (date)	Y	N	DK	1	2	3	4	5
Years of education missed as of (date)	Y	N	DK	1	2	3	4	5
Family income status (e.g., AFDC, participation in National School Lunch Program)	Y	N	DK	1	2	3	4	5
Public or private school	Y	N	DK	1	2	3	4	5
School grade level	Y	N	DK	1	2	3	4	5
B. Service Data								
Dates entered and exited project	Y	N	DK	1	2	3	4	5
Other instructional settings (e.g., mainstream class, Chapter 1 program, special education, migrant program, ESL program)	Y	N	DK	1	2	3	4	5
Dates entered and exited other settings	Y	N	DK	1	2	3	4	5

	Information Available?			If Not Available, How Accessible?				
	Yes	No	Don't Know	Very Difficult				Very Easy
C. Non-Test Data								
Years retained in grade as of (date)	Y	N	DK	1	2	3	4	5
Whether or not student dropped out	Y	N	DK	1	2	3	4	5
Days absent from project classroom as of (date) and total school days possible as of that date	Y	N	DK	1	2	3	4	5
Date of special education referral or placement	Y	N	DK	1	2	3	4	5
Date of placement in gifted and talented progra	Y	N	DK	1	2	3	4	5
Enrollment in postsecondary education institution	Y	N	DK	1	2	3	4	5
Setting from which entered project	Y	N	DK	1	2	3	4	5
Grade point average prior to entry in project	Y	N	DK	1	2	3	4	5
Setting to which exited from project	Y	N	DK	1	2	3	4	5
Grade point average after exit	Y	N	DK	1	2	3	4	5
D. Former Project Students								
Standardized achievement test data	Y	N	DK	1	2	3	4	5
Classroom performances (e.g., grade)	Y	N	DK	1	2	3	4	5
II. INSTRUCTION								
Academic and non-academic subjects taught in the project	Y	N	DK	1	2	3	4	5
Method of service provision (e.g., tutoring, classroom instruction)	Y	N	DK	1	2	3	4	5
The following variables apply to each class or subject matter offered by project:								
English and L1 materials and how they are used	Y	N	DK	1	2	3	4	5
Percents of instruction conducted in L1 and in English	Y	N	DK	1	2	3	4	5

	Information Available?			If Not Available, How Accessible?				
	Yes	No	Don't Know	Very Difficult				Very Easy
Pattern of classroom language use (e.g., concurrent, preview/review, or alternating)	Y	N	DK	1	2	3	4	5
Critical teaching strategies including instruction and management procedures (e.g., use of active teaching behaviors, use of content from students' home culture)	Y	N	DK	1	2	3	4	5
Locations of instruction (e.g., mainstream classroom, special bilingual classroom, newcomer center, resource room)	Y	N	DK	1	2	3	4	5
Student groupings (e.g., LEP students are grouped by English proficiency, all LEP students are together all day)	Y	N	DK	1	2	3	4	5
Student/teacher ratio, student/aide ratio per class	Y	N	DK	1	2	3	4	5
Variability of students within a classroom (e.g., age, grade, educational attainment, home language, language abilities)	Y	N	DK	1	2	3	4	5
Function of aides in classroom (% instruction, % management)	Y	N	DK	1	2	3	4	5
Form of instructional group (e.g., whole class, tutorial)	Y	N	DK	1	2	3	4	5
For each subject area taught, minutes per week devoted to specific activities (e.g., math workbooks, English spelling, language lab)	Y	N	DK	1	2	3	4	5

III. STAFF

Academic preparation of each staff member, including:								
• level of education,	Y	N	DK	1	2	3	4	5
• academic preparation in field of education,	Y	N	DK	1	2	3	4	5
• credentials and certificates, and	Y	N	DK	1	2	3	4	5
• academic preparation in field of bilingual/bicultural education	Y	N	DK	1	2	3	4	5

	Information Available?			If Not Available, How Accessible?				
	Yes	No	Don't Know	Very Difficult				Very Easy
Teaching experience of each staff member, including:								
• years of experience in monoligual classrooms, and	Y	N	DK	1	2	3	4	5
• in bilingual classrooms	Y	N	DK	1	2	3	4	5
Language abilities of each staff member, including languages understood and spoken and degree of fluency in:								
• English	Y	N	DK	1	2	3	4	5
• L1	Y	N	DK	1	2	3	4	5

Please answer the following questions: Date:_____

1. What is your position?_____

2. Are you involved in: data collection? Yes___ No___
   data analysis? Yes___ No___
   reporting? Yes___ No___

3. How many students are currently enrolled in your Title VII project?_____

4. What grade levels does your project serve?_____

5. What year is your project in? 1st___ 2nd___ 3rd___ 4th___ 5th___

6. How many students (approximately) are enrolled in your district?_____

7. Are your project data stored in computer? Yes___ No___

8. How are your project data analyzed? by hand? Yes___ No___
   by computer? Yes___ No___
   both? Yes___ No___

9. In your district, do you have a testing/evaluation office which: (check)
   has a centralized computer system? Yes___ No___
   provides you with some student information? Yes___ No___
   assists you in your data management? Yes___ No___
   assists you in your testing? Yes___ NO___

10. (Optional) What is your district?_____

RMC RESEARCH CORPORATION
2570 WEST EL CAMINO REAL
MOUNTAIN VIEW, CALIFORNIA 94040
TELEPHONE: (415) 941-9550

May 19, 1987

Sarah Clayton
Director of Categorical Programs
Valley Center Union School District
28751 Cole Grade Road
Valley Center, CA 92082

Dear Sarah:

On April 13, I sent you a letter and a copy of the data collection and reporting survey. This is a follow-up letter, which I hope will provide you with some guidance in preparing for the June meeting.

As you know, the June meeting is a 2-1/2 day affair. A tentative agenda has been set. During the first day, each of the nine pilot sites will be asked to make a half-hour presentation. The purpose or the presentation is to describe how various recommendations have been carried out in each project site. The contents of your presentation should include the following:

- A very brief overview of your project (title, grade levels, language(s), staff, academic or nonacademic subjects offered, number of schools, number of students)
- For all the recommendations implemented, activities undertaken to implement each (e.g., meetings classroom visitations, and so on)
- A brief description of the products resulting from implementing the recommendations (e.g., revised objectives, evaluation questions, classroom observation form, RGI, and so on)
- Problems and solutions associated with each recommendation implemented

On the second day, we will go over the User's Guide, chapter by chapter as a group, at which time I expect each site to provide us with feedback on the content (recommendations, examples, tables), format and writing. For the recommendations that you have not implemented we would like to know why, what's needed to implement them, and how sound are they. Also, what you would like to have us include that is not in the guide. It is important that you review the Guide and come up with some consensus that represents your site. I have included in this letter a list of the recommendations that are embedded in the Guide.

The third day is a day of rest for you. It will only be a half day of meetings, at which time we will be doing the talking.

I'll call you after you have received this letter.

Sincerely,

Tony Lam
Project Director

TL:mb
Enclosure

25 March, 1987

Cheryl Crawley
Salem-Keizer Public Schools
1309 Ferry Street SE
P.O. Box 12024
Salem, OR 97309

Dear Cheryl:

This is to inform you, more or less officially, that the dates for our end-of-year meeting for the "Bilingual Models" study will be June 17, 18, and 19, 1987. The meeting will be in Washington, DC.

You may bring up to three persons (including yourself) at SRA's expense. We would like you to come Tuesday evening, the 16th, so you will be ready to go by 8:30 a.m. on the 17th. We expect to break up by noon on Friday the 19th.

We have not yet selected a hotel or finalized the agenda for the meeting. This is just an "early warning" notification. We will contact you again as soon as we have more details. In the interim, please think about travel arrangements. We will want to make reservations and buy airplane tickets for your group by mid-May.

Please call me if you have any questions.

Cordially,

G. Kasten Tallmadge

GKT:TH

Letter sent to all sites

BILINGUAL EDUCATION EVALUATION SYSTEM (BEES)

WORKSHOP AGENDA

San Francisco Airport Marriott Hotel
October 7-9, 1986

Tuesday, October 7

Topic	Presenter	Time
A. Introduction		
• Welcome and self introduction	Kast	8:30 - 9:00
• Project background	Kast	9:00 - 9:15
• Purpose and guidelines of field test	Tony	
• Workshop goals and agenda/ participant expectations	Tony	9:15 - 9:30
• Pretest	Staff	9:30 - 9:45
B R E A K		9:45 - 10:00
• Overview of BEES	Tony	10:00 - 10:15
B. Evaluation Focusing	Tony	10:15 - 11:15
C. Evaluation Planning	Tony	11:15 - 11:45
L U N C H		11:45 - 1:00
D. Process Evaluation		
• What data to collect?	Tony	1:00 - 2:00
• What instrument to use?	Tony	2:00 - 3:00
• How to prepare for data collection?	Tony	3:00 - 3:15
B R E A K		3:15 - 3:30
• An example	Hector	3:30 - 4:35
E. Review and questions	Tony	4:35 - 4:50
F. Posttest	Staff	4:50 - 5:00

Wednesday, October 8 (Presenter: Kast)

Topic	Time
A. Pretest	9:00 - 9:10
B. Basic Strategies for Outcome Evaluation	9:10 - 10:00
C. Measurement Error and the Regression Effect	10:15 - 10:30
B R E A K	10:30 - 10:45
D. Selecting an Achievement Test	10:45 - 11:15
E. Other Achievement-Related Indicators	11:15 - 12:00
L U N C H	12:00 - 1:15
F. Test Administration and Scoring	1:15 - 2:00
G. Calculating Growth Indices	2:00 - 2:45
H. Review and questions	2:45 - 3:15
B R E A K	3:15 - 3:30
I. Quasi-Experimental Designs	3:30 - 4:45
J. Posttest	4:45 - 5:00

Thursday, October 9

Topic	Presenter	Time
A. Data Processing and Analysis	Kast/Tony	9:00 - 10:30
B. Integration and Interpretation of Results		
B R E A K		10:30 - 10:45
C. Evaluation Reporting	Kast/Tony	10:45 - 11:45
L U N C H		11:45 - 1:00
D. Overall review and questions	Kast/Tony	1:00 - 2:00
E. Open Discussion: Users' Guide	All	2:00 - 3:15
B R E A K		3:15 - 3:30
F. More Open Discussion: Field Test	All	3:30 - 4:00
G. Wrapup		4:00 - 4:15
H. Workshop Evaluation		4:15 - 4:25
I. Individual Questions/Farewell		4:25 - 5:00

AGENDA FOR MEETING

Washington, DC June 17-19, 1987

Wednesday, June 17

Topic	Presenter	Time
A. Introduction		
- Welcome	Kast/Jim	8:30 - 9:00
- Overview of Field Testing	Tony	9:00 - 9:15
- Overview of Meeting Agenda	Tony	9:15 - 9:30
B. Field Site Staff Presentation		
- Salem, OR		9:30 - 10:00
- Polk County, FL		10:00 - 10:30
BREAK		
- Tucson, AZ		11:00 - 11:30
- San Jose, CA		11:30 - 12:00
LUNCH		
- Valley View, TX		1:30 - 2:00
- Valley Center, CA		2:00 - 2:30
BREAK		
- Community School District #10, NY		2:45 - 3:15
- Community School District #5, NY		3:15 - 3:45
BREAK		
- Tyler, TX		4:00 - 4:30

Thursday, June 18

C. Group Discussion: Review Contents of Users' Guide (Vol. 1)		
- Introduction	Tony	8:30 - 8:4_
- Chapters 1, 2, 3		8:45 - 10:15
BREAK		
- Chapter 4, 5		10:30 - 12:00
LUNCH		
- Chapters 6, 7		1:30 - 3:00
BREAK		
-Chapters 8, 9, 10		3:15 - 4:45
-Overall Evaluation of the BEES		4:45 - 5:15

Friday, June 19

Topic	Presenter	Time
D. Summary of Feedback from Field Testing	Tony	9:00 - 10:00
E. Summary Comments from Field Staff		10:00 - 10:45
BREAK		
F. Discussion of Phase 4 Activities	Kast	11:00 - 11:30
G. Wrapup	Tony	11:30 - 12:00

APPENDIX D

Questionnaires Administered During June Meeting

STUDENT, INSTRUCTION, AND STAFF INFORMATION

I. STUDENT INFORMATION	How important is information for: Estimating Program Effectiveness (Least Important 1 2 3 4 5 Most Important)	Improving Program Performance (Least Important 1 2 3 4 5 Most Important)
A. Background Information		
Date of birth	1 2 3 4 5	1 2 3 4 5
Place of birth	1 2 3 4 5	1 2 3 4 5
Ethnicity/language group	1 2 3 4 5	1 2 3 4 5
Date began U.S. residence	1 2 3 4 5	1 2 3 4 5
Language used in home	1 2 3 4 5	1 2 3 4 5
Fluency in English as of	1 2 3 4 5	1 2 3 4 5
Fluency in native language (L1) if other than English as of	1 2 3 4 5	1 2 3 4 5
Years of education completed as of	1 2 3 4 5	1 2 3 4 5
Years of education missed as of	1 2 3 4 5	1 2 3 4 5
Family income status (e.g., AFDC, participation in National School Lunch Program)	1 2 3 4 5	1 2 3 4 5
Public or private school	1 2 3 4 5	1 2 3 4 5
School grade level	1 2 3 4 5	1 2 3 4 5
Sex	1 2 3 4 5	1 2 3 4 5
B. Service Data		
Dates entered and exited project	1 2 3 4 5	1 2 3 4 5
Other instructional settings (e.g., mainstream class, Chapter 1 program, special education, migrant program, ESL program)	1 2 3 4 5	1 2 3 4 5
Dates entered and exited other settings	1 2 3 4 5	1 2 3 4 5

	How important is information for:									
	Estimating Program Effectiveness					Improving Program Performance				
	Least Important				Most Important	Least Important				Most Important
C. Non-Test Data										
Years retained in grades as of	1	2	3	4	5	1	2	3	4	5
Whether or not student dropped out	1	2	3	4	5	1	2	3	4	5
Days absent from project classroom as of -- and total school days possible as of that date	1	2	3	4	5	1	2	3	4	5
Date of special education referral or placement	1	2	3	4	5	1	2	3	4	5
Date of placement in gifted and talented program	1	2	3	4	5	1	2	3	4	5
Enrollment in postsecondary education institution	1	2	3	4	5	1	2	3	4	5
Setting from which entered project	1	2	3	4	5	1	2	3	4	5
Grade point average prior to entry in project	1	2	3	4	5	1	2	3	4	5
Setting to which exited from project	1	2	3	4	5	1	2	3	4	5
Grade point average after exit	1	2	3	4	5	1	2	3	4	5
D. Former Project Students										
Standardized achievement test data	1	2	3	4	5	1	2	3	4	5
Classroom performances (e.g., grade)	1	2	3	4	5	1	2	3	4	5
II. INSTRUCTION										
Academic and non-academic subjects taught in the project	1	2	3	4	5	1	2	3	4	5
Method of service provision (e.g., tutoring, classroom instruction)	1	2	3	4	5	1	2	3	4	5

The following variables apply to each class or subject matter offered by project:

	How important is information for: Estimating Program Effectiveness (Least Important 1 2 3 4 5 Most Important)	Improving Program Performance (Least Important 1 2 3 4 5 Most Important)
English and L1 materials and how they are used	1 2 3 4 5	1 2 3 4 5
Percents of instruction conducted in L1 and in English	1 2 3 4 5	1 2 3 4 5
Pattern of classroom language use (e.g., concurrent, preview/review, or alternating)	1 2 3 4 5	1 2 3 4 5
Critical teaching strategies including instruction and management procedures (e.g., use of active teaching behaviors, use of content from students' home culture)	1 2 3 4 5	1 2 3 4 5
Locations of instruction (e.g., mainstream classroom, special bilingual classroom, newcomer center, resource room)	1 2 3 4 5	1 2 3 4 5
Student groupings (e.g., LEP students are grouped by English proficiency, all LEP students are together all day)	1 2 3 4 5	1 2 3 4 5
Student/teacher ratio, student/aide ratio per class	1 2 3 4 5	1 2 3 4 5
Variability of students within a classroom (e.g., age, grade, educational attainment, home language, language abilities)	1 2 3 4 5	1 2 3 4 5
Function of aides in classroom (% instruction, % management)	1 2 3 4 5	1 2 3 4 5
Form of instructional group (e.g., whole class, tutorial)	1 2 3 4 5	1 2 3 4 5
For each subject area taught, minutes per week devoted to specific activities (e.g., math workbooks, English spelling, language lab)	1 2 3 4 5	1 2 3 4 5

III. STAFF

	How important is information for: Estimating Program Effectiveness (Least Important 1 2 3 4 5 Most Important)	Improving Program Performance (Least Important 1 2 3 4 5 Most Important)
Academic preparation of each staff member, including:		
• level of education,	1 2 3 4 5	1 2 3 4 5
• academic preparation in field of education,	1 2 3 4 5	1 2 3 4 5
• credentials and certificates, and	1 2 3 4 5	1 2 3 4 5
• academic preparation in field of bilingual/bicultural education	1 2 3 4 5	1 2 3 4 5
Teaching experience of each staff member, including:		
• years of experience in monolingual classrooms, and	1 2 3 4 5	1 2 3 4 5
• in bilingual classrooms	1 2 3 4 5	1 2 3 4 5
Language abilities of each staff member, including languages understood and spoken and degree of fluency in:		
• English	1 2 3 4 5	1 2 3 4 5
• L1	1 2 3 4 5	1 2 3 4 5

PROJECT BACKGROUND INFORMATION

Community Characteristics

1. **Percent of community residents below federal poverty line.**

   - How accessible is this information? — Not Accessible 1 2 3 4 5 Easily Accessible
   - How useful is this data for providing background information about project? — Not Useful 1 2 3 4 5 Very Useful

2. **Unemployment rate.**

   - How accessible is this information? — Not Accessible 1 2 3 4 5 Easily Accessible
   - How useful is this data for providing background information about project? — Not Useful 1 2 3 4 5 Very Useful

3. **Workforce composition (blue collar, white collar, professional).**

   - How accessible is this information? — Not Accessible 1 2 3 4 5 Easily Accessible
   - How useful is this data for providing background information about project? — Not Useful 1 2 3 4 5 Very Useful

4. **Ethnic makeup of community (Asian, Black, Hispanic, Other Minority, Non-Hispanic Caucasian).**

   - How accessible is this information? — Not Accessible 1 2 3 4 5 Easily Accessible
   - How useful is this data for providing background information about project? — Not Useful 1 2 3 4 5 Very Useful

5. **Average length of residence in community.**

   - How accessible is this information? — Not Accessible 1 2 3 4 5 Easily Accessible
   - How useful is this data for providing background information about project? — Not Useful 1 2 3 4 5 Very Useful

6. Percent of single-parent families.

	Not Accessible				Easily Accessible
o How accessible is this information?	1	2	3	4	5

	Not Useful				Very Useful
o How useful is this data for providing background information about project?	1	2	3	4	5

7. Educational level of adult community residents (no high school, some high school, high school graduate, some college, college graduate).

	Not Accessible				Easily Accessible
o How accessible is this information?	1	2	3	4	5

	Not Useful				Very Useful
o How useful is this data for providing background information about project?	1	2	3	4	5

8. Community attitude toward bilingual education (generally positive, neutral or mixed, generally negative, don't know).

	Not Accessible				Easily Accessible
o How accessible is this information?	1	2	3	4	5

	Not Useful				Very Useful
o How useful is this data for providing background information about project?	1	2	3	4	5

9. Others: _____

_____

_____ _____

District Demographics

1. Approximate peak student enrollment.

	Not Accessible				Easily Accessible
o How accessible is this information?	1	2	3	4	5

	Not Useful				Very Useful
o How useful is this data for providing background information about project?	1	2	3	4	5

2. Number of schools in district.

	Not Accessible				Easily Accessible
o How accessible is this information?	1	2	3	4	5

	Not Useful				Very Useful
o How useful is this data for providing background information about project?	1	2	3	4	5

3. The school district is characterized as (rural, suburban, urban, inner city).

	Not Accessible				Easily Accessible
o How accessible is this information?	1	2	3	4	5
	Not Useful				Very Useful
o How useful is this data for providing background information about project?	1	2	3	4	5

4. Number of classrooms in bilingual program in general.

	Not Accessible				Easily Accessible
o How accessible is this information?	1	2	3	4	5
	Not Useful				Very Useful
o How useful is this data for providing background information about project?	1	2	3	4	5

5. Percent of students eligible for free or reduced-price lunch.

	Not Accessible				Easily Accessible
o How accessible is this information?	1	2	3	4	5
	Not Useful				Very Useful
o How useful is this data for providing background information about project?	1	2	3	4	5

6. Ethnic makeup of student body (Asian, Black, Hispanic, Other Minority, Caucasian).

	Not Accessible				Easily Accessible
o How accessible is this information?	1	2	3	4	5
	Not Useful				Very Useful
o How useful is this data for providing background information about project?	1	2	3	4	5

7. Others: _____

_____

_____

School Characteristics

1. Student enrollment.

	Not Accessible				Easily Accessible
o How accessible is this information?	1	2	3	4	5
	Not Useful				Very Useful
o How useful is this data for providing background information about project?	1	2	3	4	5

2. Percent of students eligible for free or reduced-price lunches.

	Not Accessible				Easily Accessible
o How accessible is this information?	1	2	3	4	5
	Not Useful				Very Useful
o How useful is this data for providing background information about project?	1	2	3	4	5

3. Ethnic makeup of student body (Asian, Black, Hispanic, Other Minority, Caucasian).

	Not Accessible				Easily Accessible
o How accessible is this information?	1	2	3	4	5
	Not Useful				Very Useful
o How useful is this data for providing background information about project?	1	2	3	4	5

4. Number of bilingual teachers (non-Title VII, Title VII).

	Not Accessible				Easily Accessible
o How accessible is this information?	1	2	3	4	5
	Not Useful				Very Useful
o How useful is this data for providing background information about project?	1	2	3	4	5

5. Number of bilingual aides (non-Title VII, Title VII).

	Not Accessible				Easily Accessible
o How accessible is this information?	1	2	3	4	5
	Not Useful				Very Useful
o How useful is this data for providing background information about project?	1	2	3	4	5

6. Number of bilingual classes (non-Title VII, Title VII).

	Not Accessible				Easily Accessible
o How accessible is this information?	1	2	3	4	5
	Not Useful				Very Useful
o How useful is this data for providing background information about project?	1	2	3	4	5

7. Others: _____

_____

_____

APPENDIX E

Tucson Unified School District Document

TITLE VII BILINGUAL SPECIAL EDUCATION

TEACHER SURVEY 1986-87

Instructions: The following items are related to your teaching or knowledge about bilingual instruction. Circle all answers that apply to each item.

1. Which of the following criteria do you use when transferring a student to English reading?

   a. Frequency with which a student uses English in speaking.

   b. Student's oral English skills must be at communicative level or beyond.

   c. Student's reading skills in Spanish are at second grade reading level.

   d. Student has been taken out of Spanish reading.

   e. Student has been given English as a second language instruction.

   f. Student reads orally in English.

   g. All of the above.

   Comments: _____

   _____

   _____

2. Which criterion do <u>you</u> think is the most important for a student to meet before transferring her/him to English Reading?

   _____

   _____

   _____

3. Which criterion do you think is <u>second</u> most important?

   _____

   _____

   _____

4. Which of the following procedures would you use in a bilingual classroom?

   a. When a student does not understand a task in English, I use Spanish to clarify the instruction.

   b. When conducting a lesson, I give directions and teach in one language, but use the other language when making informal comments to the student.

   c. Picture clues, gestures or other non-language techniques are used to clarify instruction rather than translating to the other language.

   d. Comments: _____

   _____

   _____

5. I use the following techniques in a Bilingual Special Education classroom for teaching content:

   a. When introducing a new concept to the total class, Spanish is used.

   b. Both Spanish and English are modified according to the levels of proficiency of the students in both languages when introducing a new concept.

   c. All of the above.

   d. Other (specify) _____

   _____

   _____

6. When teaching English reading; if a student does not understand a passage and asks a question in Spanish,

   a. I answer in Spanish.

   b. I answer in English.

   c. Sometimes I answer in Spanish and other times in English.

   d. I make the student ask again in English.

   e. I use gesture, picture clues, etc. to clarify.

7. When teaching Spanish or English reading,

   a. I use the preview/review method.

   b. I use the concurrent translation method.

   c. I use the phonics method.

   d. I use the whole language approach.

   e. It depends (specify)_____

_____

_____

8. The key approach to maintaining the Spanish language in reading, after transferring a student to English reading is:

_____

_____

_____

9. Engligh as a second language skills are <u>best</u> acquired if:

   a. The focus of teaching is on the <u>content</u> of students' speech rather than on correct grammer.

   b. The teacher uses only English.

   c. The teacher corrects errors in students' speech.

   d. The student is allowed to use Spanish.

   e. By intermediate age, students are interacting verbally in English.

What other factors or techniques do you thirk are important in promoting English language development?

_____

_____

_____

_____

_____

_____

APPENDIX F

Salem-Keizer Public Schools Document

SALEM-KEIZER PUBLIC SCHOOLS

Bilingual Education Program
Observation Form

Teacher _____ Observer _____ Date _____ Time _____

Aide _____ Other _____ Large Group _____ Small Groups _____

	-1-	-2-	-3-	-4-	-5-
Number of Students					
Content Area*					
Methodology+					
Appropriate (1 - 5)					
Language of Instruction					
Appropriate (1 - 5)					
Appropriateness of ESL for Students (1 - 5)					
Appropriateness of Language Use by Instructors (1 - 5)					
Students' Time on Task (1 - 5)					

Ratings: (1 = Low; 5 = High)

* **Content Areas:**

ESL
Math
Art
Language Arts
Science
Reading
Music
Social Studies
Physical Coordination
Perceptual Memory

\+ **Methodologies:**

Simplified Teacher Speech
Modeling and Gestures
Comprehension Checks
Academic Language Development
- Clustering and Webbing
- Schema Building

Cooperative Learning
TPR
Pre-teaching Vocabulary
...development

APPENDIX G

Program Documents from Valley Center, California

Valley Center

Program Goals and Objectives

1) To assist limited English proficient children to achieve English language proficiency.

To develop the child's primary oral language (Spanish).
1.1 By June of each program year, project participants that score at or below Level 4 on the BSM in Spanish and who have been in the program for nine months, will advance an average of one level in Spanish fluency as indicated by pre/post test scores on the BSM test.

1.2 After being in the project for a minimum of nine months participants that score at or below Level 4 on the BSM in English who are receiving English as a Second Language (ESL) instruction will have on an average advanced one level in English fluency as indicated by pre/post test scores on the Bilingual Syntax Measure.

2) To raise the academic achievement level of limited English proficient children.

2.1 K-8 project participants who receive reading instruction in Spanish and who have been in the program a minimum of nine months will gain on an average 1.2 months growth per month of instruction as measured by scores on the Reading and Math subtests of the Comprehensive Tests of Basic Skills (CTBS-Espanol).

2.2 LEP students in grades 2-8, who have been in the program since kinder or a minimum of two years will have gained on an average 1.2 months per month of instruction in Reading, Language and Mathematics as measured by the Comprehensive Tests of Basic Skills in English.

3) To develop Spanish as a Second Language for children whose primary language is English.

3.1 On an average, participants who receive Spanish as a second language (SSL) instruction will at the end of the first year pass 5%, the second year 15%, the third year 25%, the fourth year 50%, the fifth year 70%, the sixth year 85%, and the seventh year 100% of the items on the district made test that measures receptive and expressive skills in the second language.

VC

3.2 Project participants who have been in the program four years will score an average of a Level 2 on the BSM Spanish. Each year these participants will continue to grow an average of 1 level on the BSM Spanish until they have reached level five.

4) To reclassify limited proficient students to fluent English proficient status.

4.1 By the end of four years in the project limited English proficient students on an average 80% will be reclassified to fluent English proficient status as determined by the District's reclassification criteria.

5) To develop students positive attitudes about themselves and others.

TO/FROM

5.1 By teaching LEP students in their primary language and then transitioning them into English the positive self esteem of the students will improve as noted on the self-concept survey.

5.2 Project participants will interact socially.

6) To develop multicultural appreciation of the cultural heritage of limited English proficient children and others.

6.1 Teachers in the bilingual program will ensure that multicultural/multiethnic curriculum is integrated throughout the curriculum. Student questionnaires will be given regarding multicultural issues to assess completion of this goal.

to IPC DEVELOPMENT

7) To continue to develop and refine each project component area.

7.1 By the end of each project year the project staff will develop or refine a minimum of one component of the five project areas (instructional program, parent and community involvement, staff development, evaluation of program, curriculum materials).

8.) To continue the training for the bilingual staff working with the LEP students.

8.1 By the end of each June project staff working with LEP students will attend a minimum of one inservice related to bilingual methods and techniques.

VC

TYPES OF STUDENTS

The following is an example of five types of students which we service in the Bilingual Program. Read through the following list and if you disagree with it please make changes which you feel are appropriate.

1.- ENGLISH ONLY students who have been in the program since kinder and who's Spanish skills are well developed.

2.- ENGLISH ONLY students who entered the program after kinder and who have poor second language skills.

3.- LIMITED ENGLISH PROFICIENT students who have been in the program since kinder and have average or above average language and math skills (L1).

4.- LIMITED ENGLISH PROFICIENT students who have been in the program since kinder but, who have poor language and/or math skills (L1).

5.- NON ENGLISH PROFICIENT students who enter the program after kinder and have no English skills.

↑ HELP!

PROBLEMS IN PROGRAM IMPLEMENTATION

1.- Limited English proficient students who enter mid-year and have little or no school experience.

2.- Great variability in the second language skills of all students.

3.- Great variability in (L1) skills and levels of maturation in limited English proficient students.

4.- English only students who are put into the program without parent request or knowledge to even out numbers in other classrooms.

5.- Limited visuals and concrete objects for teaching subject content for mixed language groups.

6.- Teachers should be able to reassign students to ED classrooms if they feel they can't handle a second language.

7.- Teachers giving tests should be the daily classroom teachers which the students know and feel comfortable with.

8.- Not all parent notices which go home are transalated into Spanish.

9.- Need a secretary or part time office clerk who speaks Spanish.

10.- Need inservices for teachers who have students in ESL stage III.

11.- Need Special Education and other special services for Spanish speaking students.

OBSERVATION

The purpose of this observation is to confirm the data obtained from the bilingual teachers as critical instructional features. The information gathered by the observation is to confirm this information.

Instructions for the observer-

A classroom will be observed once a week for two weeks. The observations will be conducted for a period of 15 minutes three times a day: at the beginning of the day, mid-morning, and late afternoon. Each observer will decide the exact time of his/her observation after having discussed the classroom schedule with each individual teacher. One of the three daily fifteen minute observations must include a five minute observation of a lunch or recess period. The observer will use the time chart in the following page to schedule his/her observations.

The observer is required to answer all questions in the instrument during the early morning observation. During the mid-morning and late afternoon observation only the questions without asterisks * are to be answered. In the last question of the observation instrument the observer is encouraged to add any information which s/he feels might supplement previous observations.

June 14, 1987

VC

Critical features of the Bilingual Program a the Valley Center Union School District

Global Critical Features	Specific Critical Features	Data Sour
Use of student's first language (L1) for core subjects	*Math, reading and language are taught in the student's home language until they develop fluency in their second language. Students are encouraged to maintain their first language. *The Bilingual Syntax Measure Test and a home language survey are used to determine what language a student will be taught in.	Observation Records
Use of student's second language (L2) for non core subjects	*Art, P.E., science, social studies, music (math) are taught in the students second language.	Observation
Mixed group instruction Students are encouraged to help one another (role modeling for second language acquisition)	*Classrooms are formed 50% Spanish speakers 50% English speakers *Students are taught alternating weeks English/Spanish giving equal status to both languages.	Observation
Multicultural teaching approach. The culture of the three major ethnic groups represented is given equal importance. Other cultures are included in the curriculum.	*Curriculum materials contain information about students' home culture.	Lesson plans Review materials
Teaching second language acquisition through the "Natural Approach".	*Use of the Total Physical Response teaching technique. Teacher uses voice entonation, gestures, manipulatives, and visuals to teach. Students learn poems, songs, dramas, and write stories using their own language. *Participation of EO students in program; they are encouraged to learn a second language.	Observation
Teaching through active participation.	*Students rotate around learning centers using hands on activities. "Cooperative Learning" techniques are encouraged.	Observation

NAME OF OBSERVER_____

DAY_____DATE_____

STARTING TIME_____ENDING TIME_____

WHICH OBSERVATION IS IT?

1st_____Early morning

2nd_____Mid-morning

3rd_____Late afternoon

NAME OF TEACHER BEING OBSERVED_____

* 1.) Type of classroom:

_____Self-contained

_____Team-teaching

* 2.) Grade level_____

3.) What subject is being taught?_____

Number of times teacher speaks to students in Spanish

Number of times teacher speaks to students in English

Number of times student answers teacher in Spanish

vC

Number of times student answers teacher in English

obj #

-VALIDATION OF OBJECTIVES-
VALLEY CENTER BILINGUAL PROGRAM

READING AND LANGUAGE OBJECTIVES _____GRADE

1 Is the objective appropriate for your grade level? yes_____ no_____, why

_____
_____
_____
_____

3 Does the objective include important developmental skills the student must master for next year. yes_____ no_____

2 Does the objective cover skills that you teach. yes_____ no_____
6-6 (If you answered no on this question skip questions D and E for this objective)

) Check the approximate number of skills taught to meet this objective ☐ 1-2 ☐ 3-6 ☐ 7-12

Check the approximate number of time spent teaching the objective during the year ☐ throughout the year ☐ 1 to 2 months ☐ 1 to 3 weeks

List other skills you teach (in reading and language) and which were not mentioned in the objectives. (This question is to be answered only once)

_____
_____
_____
_____
_____
_____

Español

C.T.B.S. TEST OBJECTIVES

OBJECTIVOS

nivel	A	B	C	1	2	3
Reconocimiento del sonido						
1. El estudiante identificará correctamente la palabra dada oralmente.		1.6k O				
2. El estudiante identificará correctamente la palabra que significa algo sobre una palabra o una oración dada oralmente.			2.6 O			
Comprensión de lectura						
3. El estudiante identificará correctamente el dibujo que enseña el significado de una oración escrita.		1.6a O				
4. El estudiante identificará correctamente la palabra que significa algo sobre un dibujo que vee.		1.6d O				
5. El estudiante identificará correctamente la palabra que mejor complete la oración escrita.			O			
6. El estudiante identificará correctamente la palabra que significa algo sobre la palabra o frase que está subrayada.				O	O	O
7. El estudiante contestará correctamente preguntas como: quién, qué, dónde, y cuándo, sobre pasajes que lee.			O	3.6	O	

G-10

C.T.B.S. TEST OBJECTIVES

nivel	A	B	C	1	2	3
Comprensión de lectura						
8. El estudiante podrá analisar sentimientos o motivos de los carácteres en un pasaje que lee.			2.5	3.1		
9. El estudiante identificará la idea principal, el intención del autor, o el humor y el tono expresado en un pasaje que lee.			2.6	3.7		
10. El estudiante sacará conclusiones del pasaje que lee.			2.8	3.9		
Computación de matemáticas						
11. El estudiante sumará números enteros.						
12. El estudiante sumará decimales o números quebrados.						
13. El estudiante restará números enteros.						
14. El estudiante restará decimales o números quebrados.						
15. El estudiante multiplicará números enteros.						
16. El estudiante multiplicará decimales o números quebrados.						
17. El estudiante dividará números enteros.						
18. El estudiante dividará decmales o números quebrados.						

C.T.B.S. TEST OBJECTIVES

nivel	A	B	C	1	2	3
Conceptos y Aplicaciones de matemáticas						
19. El estudiante demonstrará que entiende la numeración.						
20. El estudiante demonstrará que entiende operaciones de números. (number sentence)						
21. El estudiante demonstrará que entiende a resolver problemas.						
22. El estudiante demonstrará que entiende la teoría de números.						
23. El estudiante demonstrará que entiende medidas, dinero, y tiempo.						
24. El estudiante demonstrará que entiende geometría.						

VC

June 14,1987

C.T.B.S. TEST OBJECTIVES - VALIDATION
First grade - READING

RECONOCIMIENTO DE SONIDO
1.- El estudiante identificara correctamente la palabra dada oralmente. #1.6b of VC OBJECTIVES

COMPRENSION DE LECTURA
3.- El estudiante identificara correctamente el dibujo que ensena el significado de una oracion escrita. #1.6c of VC OBJECTIVES

4.- El estudiante identificara correctamente la palabra que significa algo sobre un dibujo que ve. #1.6d of VC OBJECTIVES

The greatest concern among first grade teachers involving these three objectives was how far into the reading sequence the objective measured.

Spanish reading in Valley Center is taught in a phonetic global manner with a continuous emphasis in oral language development. The child is first taught the vowel sounds and their combinations. The next step is to introduce consonant sounds one at a time and from the very beginning whole sentences are stressed. These sentences contain only the vowels and the consonant(s) which have been introduced. A specific sequence is followed when introducing consonants the first seven letters being M,S,P,T,L,D,and N.
The above objectives do not state if only certain consonants are used or if the child is expected to know the entire alphabet. If the child is expected to read words containing all the letters of the alphabet only about 50% of the Valley Center first graders (who read in Spanish) would attain these objectives.

VC

C.T.B.S. TEST OBJECTIVES - VALIDATION
Second grade - READING

RECONOCIMIENTO DEL SONIDO
2.- El estudiante identificara correctamente la palabra que significa algo sobre una palabra o una oracion dada oralmente. #2.3b VC OBJECTIVES

COMPRENSION DE LECTURA
5.- El estudiante identificara correctamente la palabra que mejor complete la oracion escrita. #2.4 VC OBJECTIVES

8.- El estudiante podra analisar sentimientos o motivos de los caracteres en un pasaje que lee. #2.5 VC OBJECTIVES

9.- El estudiante identificara la idea principal, la intencion del autor, o el humor y el tono expresado en un pasaje que lee. #2.6 VC OBJECTIVES

10.- El estudiante sacara conclusiones del pasaje que lee. #2.8 VC OBJECTIVES

Second grade teachers agreed that all of the above reading objectives were appropriate for their grade level. They agreed that the objectives included important developmental skills the students needed to master for the following year. They all taught a large number of skills throughout the year to meet these objectives.

VC.

C.T.B.S. TEST OBJECTIVES - VALIDATION
Third grade - READING

COMPRENSION DE LECTURA

6.- El estudiante identificara correctamente la palabra que significa algo sobre la palabra o frase que esta subrayada. #3.5a & b VC OBJECTIVES

7.- El estudiante contestara correctamente preguntas como: quien, que, donde, y cuando, sobre pasajes que lee. #3.6 VC OBJECTIVES

8.- El estudiante podra analisar sentimientos o motivos de los caracteres en un pasaje que lee. #3.12 VC OBJECTIVES

9.- El estudiante identificara la idea principal, la intencion del autor, o el humor y el tono expresado en un pasaje que lee. #3.7 VC OBJECTIVES

10.- El estudiante sacara conclusiones del pasaje que lee. #3.9 VC OBJECTIVES

Third grade teachers agreed that all of the above reading objectives were appropriate for their grade level. They agreed that the objectives included important developmental skills the students needed to master for the following year. They all taught a large number of skills throughout the year to meet these objectives.

VC

EVALUATION QUESTIONS

1.) Are students who have been in the bilingual program four years doing their core subject in their second language?

2.) During unstructured play do English and Spanish speakers interact socially?

3.) Does the teachers' attitude towards Bilingual Education affect the students perception about learning a second language?

4.) Is the success of the program due to the methods teachers use or more to the fact that we are good teachers and care about the students?

5.) Do Spanish component teachers use more English than they should?

6.) Is the success in ESL due to good ESL instruction or exposure to English in the environment (on playground, bus, TV, etc.)

7.) Is the program meeting the needs of non English speakers who enter the program after second grade?

8.) Are students in bilingual classrooms truely more culturally aware?

9.) How do the students in the program feel about being involved in the program?

10.) Have the changes that have been made in the program's component areas (parental involvement, staff training, evaluation, instruction, and curriculum and materials) been useful?

11.) Are students who have participated in the project graduating from high school?

12.) Are students who are participating in the team teaching approach learning Spanish faster than those in the self contained?

13.) Are fewer students being retained?

APPENDIX H

San Jose Unified School District Document

TO: Yolanda

~~Obervation Issues~~ - CALIBRE/VOICES CLASSROOM OBS. FORM

1. Amount of time with computer assisted instruction: _____
   teacher: _____ aide: _____
2. Language used by instructional aide: _____

SAMPLE FORM

Instructions:

I. Background

A. Name of aide/~~teacher~~: _____ teacher: _____
B. School: _____
C. Subject area: _____ grade
D. No. of students observed: _____
E. Type of bilingual methodology employed: _____
F. Type of learning environment: ~~classroom~~ whole class ___ Small group ___ Individual ___

Comments on environment, groupings etc. _____

II. Observation Data (teacher/aide) _____

A. Language use (estimate) _____

Teacher: 1. L_1 ___% Aide L_1 ___% Student L_1 ___%
2. L_2 ___% L_2 ___% L_2 ___%
3. other (please specify) _____

Comments _____

B. Time on learning task

1. with teacher: _____ minutes
2. with aide: _____ minutes
3. with computer assisted program: _____ minutes
4. with Resource teacher _____ "
5. other _____ "

C. Student Participation

1. ~~number of student responses:~~ _____
2. ~~type of student responses:~~ _____
3. are students involved in relevant follow-up activities? Yes ___ No ___

D. Integration of Cultural ~~events~~ background of students: _____

1. use of relevant examples: _____
2. ~~other:~~ use of ~~appropriate~~ materials culturally relevant _____

E. Comments: _____

E. Learning Environment

1. Is the person's language (L_1) used in the classroom? Yes___ No___
2. Are culturally relevant materials employed? Yes___ No___
3. Other (please specify)

III. Observer Information

1. Name:_____
2. Date:_____
3. Duration of Observation:_____
4. Comments:_____

_____

_____

_____

Observer's Signature: _____

Please submit this completed form to:

APPENDIX I

New York Community School District #10 Document

COMMUNITY SCHOOL DISTRICT TEN
BOARD OF EDUCATION OF THE CITY OF NEW YORK
P.S. 95 · 3961 HILLMAN AVENUE
BRONX, N.Y. 10463

TELEPHONE 220-6300

FRED GOLDBERG
COMMUNITY SUPERINTENDENT

SANDRA LERNER
DEPUTY COMMUNITY SUPERINTENDENT

Suggested Criteria for English As A Second Language Classroom Observations

A. The Lesson

1. Is the language used appropriate for student's level of proficiency?
2. Are words and forms used in context? (In questions, sentences, and situations)
3. Are the aims of the various major sections of the lesson (review of the previous work, boardwork, new presentation) clear?
4. Are the various major sections of the lesson motivated?
5. Is the presentation of the new work well organized?
6. Is the presentation of the new work clear?
7. Are there checks made on the comprehension of the new presentation?
8. Is a sufficient amount of time devoted to the new presentation? (e.g. 15-20 min.)
9. Is a major part of the lesson devoted to oral activities?
10. Is there a variety of activities? (dialogue, question and answer, singing, drills, audio-visual, etc.)
11. Does the teacher use devices for heightening interest in the lesson (personalization, situational approach, visual supports, etc.)?
12. Is the pace of the lesson appropriate to the activity and to the level of the students?
13. Are cultural elements introduced as part of the lesson?
14. Does the lesson have an over-all structure? (e.g. warm up, review, new material, application, summary, assignment, etc.)

B. Questioning and Student Participation

1. Are most of the students involved in the lesson?
2. Are questions used for a variety of purposes (eliciting, probing of comprehension, reenforcing, personalizing)?
3. Are questions clear?
4. Are questions individualized?
5. In calling on students, is there a balance between volunteers and non-volunteers?
6. What is the quality of responses?

C. Routines

1. Are directions clear?
2. Is the time of the period used economically and appropriately?
3. Are there transitions to new activities?
4. Does the teacher have an effective manner of starting the lesson?
5. Does the teacher have an effective manner of distributing and collecting materials?

D. Boardwork and Use of Aids

1. Is there proper utilization of the boards? (assignment, new vocabulary, correction of assigned material, dictation, drills, fill-ins, completions, summaries)
2. Is most of the writing on the boards done by students?
3. Does the teacher use visual supports effectively?
4. Is the illustrative material appropriate and integrated with the lesson?

E. The Teacher

1. Is the teacher prepared for the lesson?
2. Does the teacher have command of basic techniques of teaching E.S.L. teaching (introduction of new material, review, boardwork, drills, reenforcement)?
3. Does the teacher have rapport with the students?
4. Does the teacher use instructional materials effectively
5. Is the teacher adapting the textbook to his/her class needs?
6. Does the teacher show personal qualities necessary for effective teaching (vitality, concern for all students, eagerness to praise)?

Rating Scale for Second Language Classroom Evaluation

The Lesson	Superior	Good	Fair	Unsatisfactory
Extent to which the second language is used in the classroom				
Functional use of the language				
Motivation				
Clarity of presentation				
Evidence of comprehension of the presentation				
Structure of the presentation				
Eliciting of information				
Use of words in context				
Adequacy of drill				
Extent of oral activities				
Over-all structure of the lesson				
Appropriateness of activities				
Variety of activities				
Pace				
Dramatic quality of the lesson				
Spontaneity				
Cultural enrichment				
Personalization				
Evidence of co-curricular activities				
Use of materials other than conventional instructional materials (newspapers, magazines, advertisements, posters, etc.)				
Outcomes				
Questioning				
Distribution of questions				
Individualization of questioning				
Clarity of questions				
Adequacy of questioning				
Proper use of questioning (eliciting, probing, reenforcing, personalizing)				

Teacher	Superior	Good	Fair	Unsatisfactory
Preparation for the lesson				
General preparation				
Command of techniques				
Linguistic skills				
Rapport with students				
Vitality				
Concern for individuals				
Independence from the textbook				
Textbook adaptation				
Students				
Participation				
Quality of responses				
Student-centered approach				
Individualization				
Initiative for self-direction and cooperation				
Development of good attitudes				
Routines				
Clarity of directions				
Efficiency in distribution of materials				
Transitions to new activities				
Economical use of time				
Start of lesson				
Boardwork				
Utilization of boards (assignment, new vocabulary, correction of assigned material, dictation, drills, fill-ins, completions, summaries)				
Correction of boardwork				
Student participation in boardwork				
Orderliness				
Use of Aids				
Appropriate use of A-V aids				
Use of teacher or student made materials				
Use of illustrative materials				

APPENDIX J

New York Community School District #5 Document

#45

June 15, 1987 - Second Revision

PROGRAM OBJECTIVES FOR TRANSITIONAL BILINGUAL EDUCATION PROJECT - GRADES K-3

LONG RANGE OBJECTIVES (THREE YEAR)

Student Outcomes:

1. By the end of the project period, 85% of the target population (those students receiving instruction for two or more years) will demonstrate higher-order thinking skills ability as measured by correctly answering 75% of the questions on a program developed and program related criterion referenced test.

2. By the end of the project period, 85% of the target population (those students receiving instruction for two or more years) will domonstrate improvide proficiency in speaking and understanding the second language (English) by obtaining scores above the 20th percentile on the Language Assessment Battery (LAB).

3. By the end of the project period, the target population (those students receiving instruction for two or more years in the Thinking Skills Program) will demonstrate a mean Relative Growth Index (RGI) of over 100 percent in reading in the native language (Spanish) as measured by the Comprehensive Test of Basic Skills, Reading Section, Spanish Version, between pretest and posttest scores.

4. By the end of the project period, the target population (those students receiving instruction for two or more years in the Thinking Skills Program) will demonstrate a mean Relative Growth Index (RGI) of over 100 percent in mathematics in the native language (Spanish) as measured by the Comprehensive Test of Basic Skills, Math Section, Spanish Version, between pretest and posttest scores.

5. By the end of the project period, the target population (those students receiving instruction for two or more years in Thinking Skills) will demonstrate a mean Relative Growth Index (RGI) of over 100 percent in mathematics in the second language (English) as measured by the Metropolitan Achievement Test, Math Section, between pretest and posttest scores.

6. By the end of the project period, the target population (those students receiving instruction for two or more years in Thinking Skills) will demonstrate a mean Relative Growth Index (RGI) of over 100 percent in reading in the second language (English) as measured by the Degrees of Reading Power, Reading Test, between pretest and posttest scores.

Process Objectives:

1. By the end of the project period, 80 percent of those staff receiving training will demonstrate the ability to effectively use instructional techniques that provoke students' thinking as evidenced by a mean rating of 4.0 or higher on a five point scale set in the form of two checklists of techniques and behaviors (one for teachers and one for aides). The data will be gathered through the use of classroom observations. Only teachers and aides receiving training for a minimum of one year and teaching or working as an aide for two years in the program will be included.

2. By the end of the project period, 80 percent of those staff receiving training will report on their ability to effectively use instructional techniques that provoke students' thinking as evidenced by a mean rating of 4.0 or higher on the Levels of Use rating system. A rating of 4.0 connotes routine use of an innovation. The data will be gathered through the use of teacher and aide interviews. Only teachers and aides receiving training for a minimum of one year and teaching for two years in the program will be included.

Parental Involvement:

1. By the end of the project period, 70 percent of target population parents will demonstrate knowledge and understanding of the role of thinking in their childrens'' education by correctly answering 75% of the questions on a program developed and program related parent questionnaire.

2. By the end of the project period, 70 percent of target population parents will have attended one or more school-related function or meeting connected with this program or their students' education.

Program Management and Implementation:

1. By the end of the project period, a program of transitional [illegible]nguel education from kindergarten through third grade using [illegible]inking skills will have been fully established as evidenced by documentation recorded in project logs.

2. By the end of the project period, a manual of lessons, classroom materials and ideas will be developed and made available to other educators.

OBSERVATION CHECKLIST

BACKGROUND INFORMATION

Observer _____

Person(s) Observed _____

(1) Name _____

Position _____

(2) Name _____

Position _____

School _____ Grade _____ Class _____

Date _____ Period _____

Subject Matter Taught _____

Number of Students on Register _____ Present _____

Length of Observation _____

CLASSROOM OBSERVATION CHECKLIST

TEACHER/AIDE

Thinking Skills Strategies

1. Does teacher pose open-ended questions?

2. Does teacher use probing techniques to elicit additional information?

3. Does teacher allow sufficient time for student to ponder question?

4. Does teacher respond to answers in a postive manner?

5. Does teacher permit and encourage questions from students?

6. Does the teacher use higher level thinking skills teaching techniques (classifying, analysis, comparison, infering, evaluation) when teaching language arts, science, social studies, and math (Based on class observations and/or students performance folders)?

Methodology

1. Do lesson plans reflect lesson taught?
2. Does the lesson have a clear aim?
3. Does teacher adjust for student ability?
4. Do slower students receive extra help?
5. Are charts, materials and visuals comprehensible?

Management

1. Do students complete work during assigned periods?
2. Are assignments checked?
3. Are assignments differentiated by student ability?
4. Do students have personal achievement records?
5. Do students share ideas and products of their work?
6. Is student discipline good?

Environment

1. Do teachers and students have sufficient space to work?
2. Is there provision for students to work in small groups?
3. Are there centers for students to work alone or in pairs?
4. Does the classroom have educationally stimulating decorations?
5. Are the learning materials well organized and easily accesible?
6. Does the classroom have visual cultural representation?
7. Is student work displayed and is it current?
8. Is there a class library?
9. In general, is the classroom environment comfortable and conducive to stimulate creativity?
10. Are there facilities for students to experiment, and make discoveries?
11. Does visitors log reflect parental involvement?

Interaction

1. Does the teacher demonstrate respect for students?
2. Do the students demonstrate trust in the teacher?
3. Do students treat each other in friendly and courteous manner?
4. Is there evidence of close collaboration between teacher and aide?
5. Is there evidence of cooperation with other staff members in the school?
6. Do students practice peer tutoring?
7. Are students free to move to designed learning centers in the classroom at will?