DOCUMENT RESUME

ED 102 817 FL 005 903

AUTHOR	Offenberg, Robert M.; And Others
TITLE	Title VII Bilingual Project "Let's Be Amigos": Evaluation of the Fourth Year, 1972-73.
INSTITUTION	Philadelphia School District, Pa.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
PUB DATE	Oct 73
GRANT	OEG-0-9-480089-3503
NOTE	104p.
EDRS PRICE	MF-$0.76 HC-$5.70 PLUS POSTAGE
DESCRIPTORS	Academic Achievement; *Biculturalism; *Bilingual Education; *Bilingual Schools; *Bilingual Students; Curriculum Evaluation; Educational Assessment; English (Second Language); Evaluation Methods; *Program Evaluation; Teaching Procedures; Testing
IDENTIFIERS	Elementary Secondary Education Act Title VII; ESEA Title VII; Philadelphia

ABSTRACT

This is a report on the fourth year of the Let's Be Amigos bilingual project. The project consists of three components: Model A provides bilingual/bicultural education to English-dominant and Spanish-dominant pupils in prekindergarten through fourth grade; Model B provides bilingual/bicultural education to Spanish-dominant pupils, first through third grade; Arriba provides Spanish-language instruction with English taught as a second language, grades three through twelve. Part 1 examines program processes and includes: general evaluation; curriculum reorganization and evaluation; and evaluations of Model B by principals of participating schools. Part 2 studies pupil performance and includes: pupil performance evaluation; description of standardized testing; and a study of the effect of participation in the Arriba program on graduation prevalence. Data analyses and samples of materials illustrate the report, and a bibliography is appended. (AM)

TITLE VII BILINGUAL PROJECT
"LET'S BE AMIGOS"
EVALUATION OF THE FOURTH YEAR
1972-1973

OFFICE OF RESEARCH
AND EVALUATION
THE SCHOOL DISTRICT OF PHILADELPHIA

TITLE VII BILINGUAL PROJECT
"LET'S BE AMIGOS"
EVALUATION OF THE FOURTH YEAR, 1972-1973

Report Prepared by:

Robert M. Offenberg
Research Associate

Judith Flint
Research Intern

Leonor Lega
Research Intern

Project No. 7404

Division of Instructional Research and Development Services

Edward K. Brown
Director

Office of Research and Evaluation

Michael H. Kean
Executive Director

I. Ezra Staples
Associate Superintendent
Curriculum and Instruction

David A. Horowitz
Associate Superintendent
School Services

Eleanor L. Sandstrom
Director
Foreign Languages Education

Thomas C. Rosica
Director
Federal Programs

Matthew W. Costanzo
Superintendent of Schools

THE SCHOOL DISTRICT OF PHILADELPHIA

October 1973

THE SCHOOL DISTRICT OF PHILADELPHIA

Superintendent of Schools
Dr. Matthew W. Costanzo

Executive Deputy Superintendent
Dr. Robert L. Poindexter

Members of the Board of Education

Mr. William Ross, President
Robert M. Sebastian, Esq., Vice President
Mr. Augustus Baxter
Mrs. Lawrence I. Boonin
Dr. Philip Davidoff
Mr. George Hutt
Mrs. Edward Oberholtzer
Mr. Arthur W. Thomas
Dr. Alec Washco, Jr.

ACKNOWLEDGMENTS

This program is supported in part by a grant of the United States Office of Education. The grant, awarded under Title VII ESEA, is number OEG-0-9-480089-3503.

Program supervisors Felicita Hanna, Ruth Montalvo, Mario Ramirez, Ramona Rodriguez, and Minerva Dean assisted with the collection of the data presented herein; Robert Reiter edited the report. Without the efforts of these persons, this report could not have been completed.

ABSTRACT

The 1972-1973 LET'S BE AMIGOS Title VII program provided bilingual-bicultural education to English-dominant and Spanish-dominant pupils in eleven schools. This was its fourth operational year. The program consisted of three educational components whose aim was to meet the unique needs of the Spanish-speaking while fostering bilingual-biculturalism among all participants.

MODEL A provided bilingual-bicultural education to all pupils, English-and Spanish-dominant, enrolled in prekindergarten through Grade 4 at the Potter-Thomas School. It used team teaching to provide instruction in both English and Spanish languages to pupils in all subject areas.

MODEL B provided bilingual-bicultural education to Spanish-dominant children in selected first-,second-, and third-grade classes. The instructional procedure was similar to that provided to the Spanish-dominant group of MODEL A. It was conducted at Miller and Ludlow Schools.

ARRIBA provided instruction in major subject areas in Spanish, and instruction in English as a second language. It served pupils in Grades 3-12 in Ludlow, Waring, and Southwark elementary schools, Penn Treaty and Stoddart-Fleisher junior high schools, and Edison, Kensington, Franklin, and Penn Senior high schools.

During the 1972-1973 school year there were two lengthy teacher strikes. Despite these two disruptions in the education of the children, major findings of the current evaluation showed that performance was generally better, and at no time worse, than preprogram performance.

Spanish reading performance of Spanish-dominant children was superior to that of preprogram base-line groups in virtually all levels and components of the program. In the Model School components, this performance was similar to that obtained in Puerto Rico by the test makers.

English language-arts performance of Spanish-dominant children was at the target in third grade and somewhat behind the target level in fourth grade (the two only levels in which it was tested).

Standardized test performance of English-dominant pupils was at least the same as, and often superior to, preprogram baselines.

Spanish reading performance of the English-dominant groups was far superior to that expected.

In the high schools, the program has been effective in greatly increasing the number of pupils who graduated school.

In addition to these findings, process-evaluation studies revealed principal's attitudes toward bilingual education, strengths and weaknesses of the curriculum development, and ways in which some program objectives needed to be modified.

SUMMARY OF THE "LET'S BE AMIGOS" PROJECT EVALUATION

The LET'S BE AMIGOS program has been in operation for four years. As is the case with any complex program, over this time span the evaluations of the program have changed; as some questions were answered with high degrees of confidence in the evaluations of the first years, they have been replaced by new questions, appropriate for a more mature program. The early evaluation results are believed to continue to be a fair representation of program outcomes, and are therefore still relevant.

This summary provides a review of the major findings in the LET'S BE AMIGOS program to date. Each study comprising this report is taken up in turn. If a study was a replication of previous evaluative research in the program, the fact is noted, as are any differences in the outcomes. Following this summary, major outcomes of previous evaluations not replicated are cited.

Study 1. General Process Evaluation

In the fourth operational year, the three components of the Bilingual Program were serving 1,834 students--1,189 in the MODEL A Bilingual School, 170 in the MODEL B Bilingual School program, and 475 in ARRIBA programs. Of these students 1,110 were judged to be Spanish-speaking, 184 were judged to be speakers of both Spanish and English, and 523 were judged to be speakers of English. (Data are lacking for the remaining few students.) The English-speaking pupils were virtually all participants in MODEL A; MODEL B and ARRIBA students were nearly all Spanish-speaking.

The MODEL A program operated in the Potter-Thomas School, prekindergarten through fourth grade, with teams of English- and Spanish-speaking teachers working with the pupils. In the early grades instruction was predominantly in the mother tongue. As pupils matured, increased contact with the second language was built into the program.

Two schools--Miller and Ludlow--had first- through third-grade classes which comprised the MODEL B component. The course of study was similar to that of MODEL A for Spanish-dominant pupils. Spanish-speaking teachers worked with classes of pupils dominant in the teacher's mother tongue. Itinerant second-language specialists worked with the pupils in their specialty--English as a second language.

Eight schools--two elementary, two junior high, and four senior high schools--comprised the ARRIBA program. This served primarily Spanish-speaking pupils who could benefit from, or wanted instruction in, their mother tongue. At all levels courses were offered in Spanish as a first language, English as a second language, social studies in Spanish, and science in Spanish. The content of the courses varied with grade level and was (where possible) similar to the English-language program studied in the Philadelphia schools. At the elementary and junior high schools pupils served by the program took all subjects offered for their grade levels. At the high school level pupils could select courses from among those offered by the program. In one high school--Kensington--some commercial courses in Spanish were added to the ARRIBA component during the 1971-1972 year and were continued in

1972-1973.

During 1972-1973 there were two uncontrollable events which possibly affected pupil performance and made it atypical: two citywide teacher strikes resulted in extensive closing of schools, and the illness and death of one of the program supervisors left a void in the program management which could not be filled completely.

Principals decided to drop English-speaking pupils from participation in the MODEL B program. Study 4 (See below) describes the principals' action in detail.

Study 2. Reorganization of MODEL A and MODEL B Micro-objectives

Micro-objectives are the individual behavioral bits which the Model School components are designed to teach. In each of the preceding years, as these components cycled upward through the grades, new and supposedly more difficult objectives were adopted for the added grade levels. By the beginning of the fourth year, it appeared that the micro-objectives were not in the optimal order for an orderly instructional program. This study shows how expert teachers reordered the micro-objectives into a continuous scale. In all subject areas, the agreement among raters was very high. Results showed that in two areas--language arts (first and second language) and social studies--the order in which these teachers felt material should be mastered was quite different from that appearing in the original sequence. This suggested that revision of the curriculum sequence was needed.

Study 3. Evaluation of Social Studies and Science Materials

Materials developed for the junior high school social studies and science curricula were evaluated through criterion-referent testing of seventh-grade pupils who worked with them. Results showed that the materials, as they were taught by the program staff, did not result in adequate performance. Careful review of the materials and the teaching procedures used was suggested.

Study 4. Principals' Perceptions of MODEL B.

MODEL B has been the most difficult component of the LET'S BE AMIGOS project to implement. This study attempted to gain insight into the difficulties by reviewing the component with four principals who have or have had the program in their schools. The major finding was that three of the principals who have worked with this component generally feel uneasy about the instruction in a language other than English, especially when the instruction involves English-speaking pupils and when the instruction is more than a temporary measure to meet the needs of children with problems.

The fourth principal dropped the program when the number of Spanish-speaking pupils fell below earlier expectations. This fourth principal has maintained some bilingual instruction supported by her own school's

regular funds to meet the needs of the few Spanish-speaking pupils on roll.

Study 5. Kindergarten and Prekindergarten Pupil Performance in MODEL A

Both English-dominant and Spanish-dominant kindergarten pupils were tested using the Philadelphia Readiness Test in the pupils' mother tongues. Both pupil groups were performing better than the preprogram base-line groups. This was a replication of performance found in the first and third years of program operation, and was superior to pupil performance in the program's second operational year.

In addition to the Philadelphia Readiness Test, a newer instrument--the Boehm Test of Basic Concepts--was tried. In the Spanish-dominant group, pupils were at the 50th percentile (Low S.E.S. norms). English-dominant pupils were at the 65th percentile; but variation in the testing procedures used might account for the observed differences.

Within the kindergarten group, a special all-day kindergarten (consisting mainly of pupils with prekindergarten experience) was held. Results with both the Philadelphia Readiness Test and the Boehm Test showed highly superior levels of performance, with Boehm average scores at the 75th percentile (Spanish-dominant) and the 70th percentile (English-dominant).

Use of part of the Boehm test in the prekindergarten suggested that there was enough variability in scores to make it useful in discriminating candidates for all-day kindergarten from those who should be enrolled in regular kindergarten.

Study 6. Standardized Testing of Pupils in Second Through Fourth Grades

This testing examined performance of program participants on norm-referenced instruments which measured reading and (in English) other skills. Results showed that in all components (except for a small group of ARRIBA pupils in fourth grade) Spanish-dominant pupil performance in Spanish was about good as that of pupils in rural Puerto Rico (mean scores ranged from the 40th to the 65th percentile). This compared with base-line performances which ranged from the 5th to the 35th percentile.

English-dominant pupil performance in English was significantly higher than the historical base-lines in the second grade, and about equal to that of the base-line groups in the third and fourth grades.

In the second language, Spanish-dominant pupils performed at expected levels in the third grade but were below the anticipated levels in the fourth grade.

Fourth-grade English-dominant pupils' performance in Spanish was surprisingly good, with pupils performing like third-grade native Spanish speakers.

Results showed that, despite the interruptions due to the strikes, Spanish-language performance in the second- through fourth-grade levels was better than anticipated and better than observed in previous years. English-language performance was about the same as that of pupils in the base line despite these disruptions.

Study 7. Standardized Testing of Reading in Junior High ARRIBA

This study replicated procedures used in the 1971-1972 school year to assess Spanish reading performance. The pupils in the seventh and eighth grades of the program were substantially above the base line, but not quite so good as in the 1971-1972 school year, possibly due to the strikes. In contrast, pupils in the ninth grade performed better than the relatively low level of pupils in the program last year.

Study 8. Effect of ARRIBA Participation on Graduation Rate

In two previous evaluations, it was found that the dropout rate of pupils in the program was reduced during the school year. This study demonstrated that the observed reductions were maintained over the years, increasing the number of pupils who graduated. The percentage of pupils graduating in 1973 was computed for Spanish-dominant pupils who had been in the program in tenth grade and those who had not. Among Spanish-dominant pupils attending the four high schools served by the program, participants were four times as likely to graduate as were pupils not participating. Citywide, pupils in the program were nearly twice as likely to graduate as were nonparticipating Spanish-dominant pupils.

Major Findings in Previous Reports

Previous reports have shown some other facts, not studied during 1972-1973:

1. Pupils mastered arithmetic and writing skills at or above levels specified in objetives on specially designed criterion tests (MODEL A).

2. Teachers believe that pupils show more adaptive classroom behaviors when in classes in which their mother tongue is the medium of instruction (MODEL A).

3. Pupils in the program tend to have more improved grades and behavior ratings than control pupils (ARRIBA).

4. Parents, teachers, and principals have generally been supportive of the program and have wished that the children continue to participate in it.

Conclusion

Overall results show that the LET'S BE AMIGOS project has improved performance of most Spanish-speaking participants in most areas examined. The project has, as a minimum, introduced English-speaking children to a second language without apparent harm, and has frequently enhanced their performance in some academic areas in English.

CONTENTS

STUDY 1

GENERAL PROCESS EVALUATION

In the first two evaluation reports prepared for the LET'S BE AMIGOS program, comment was made on individual process objectives. Beginning with the third year, process evaluation has focused on specific issues which arose during the school year and required attention. This year's evaluation information is again being presented in three ways as it was last year. Data gathered in time for the continuation proposal, appeared in that document (See the LET'S BE AMIGOS Continuation Proposal, 197 -1974); data not available for the continuation proposal or which needed to be updated at the end of the year appear in this study. In addition, questions which were examined extensively are presented as separate studies in this report (Studies 2, 3, and 4).

Description of the Components

MODEL A. The MODEL A program continued to operate at the Potter-Thomas School, where it encompassed prekindergarten through fourth-grade classes, the last of these being new during the fourth operational year.

The MODEL A program plan was to team-teach, with native English-speaking teachers and native Spanish-speaking teachers (nearly all of whom had at least working knowledge of both languages) teaching pairs of classes. Pupils in each pair of classes met in linguistically mixed or homogeneous groups, with most instruction in homogeneous, mother-tongue groups in the lower grades and greater amounts of second-language contact in both homogeneous and mixed groups in the upper grades.

During the typical school day, pupils moved between teachers in the team, assuring that each child had some experiences in this mother tongue and his second language and in ethnically homogeneous and ethnically diverse groups. The curriculum specified materials to be used, global pupil-performance goals to be obtained, and micro-objectives describing tasks that pupils should have been able to perform.

As the pupils progressed through the grades, their contact with the second language was to increase from 10% of the school day at the beginning of the prekindergarten, kindergarten, and first grade to one-half day in each language in third and fourth grades. The prekindergarten and regular kindergarten levels of the program met for a half-day of school; the other levels had full-day programs.

Beginning in the second operational year, a special all-day kindergarten program was developed in which pupils began reading preprimer materials in their mother tongue. This plan was maintained in 1972-1973, and alumni of the 1971-1972 all-day kindergarten were enrolled in a specially enriched first-grade curriculum.

An innovation in the MODEL A program developed in 1972-1973 was the initiation of out-of-grade instruction, especially in the second language. As the program cycled upward, pupils who had not had bilingual educational experience entered the classes. These pupils were assigned to the teachers working at the appropriate instructional level in their first and second languages even if most pupils in the teacher's class were at another grade level.

MODEL B. This program sought to bring a bilingual school experience to smaller groups of English and Spanish-speaking pupils within a school which was not reorganized as a bilingual program. According to the original plan at each site, at least two classes in one grade level--one class consisting of Spanish-dominant children and one consisting of English-dominant children--were to participate in the program. Classes received second-language instruction from specialists in those subject areas. The classroom teachers had to be native speakers of the language in which they instructed the students.

This program was in operation in the Miller and Ludlow elementary schools. The MODEL B program was withdrawn from the Bethune-Wright school because there were fewer Spanish-speaking pupils in the school than anticipated.

In 1972-1973, the MODEL B program did not operate in accordance with the above plan, because principals withdrew their English-speaking children from the program. Detailed discussion of the principals' actions is included in Study 4.

The instructional program and objectives of MODEL B paralleled those of MODEL A, except that new curricular materials prepared in Miami by the Spanish Curriculum Development Center were used extensively in MODEL B.

ARRIBA. The ARRIBA program continued during the third operational year to offer grade-level-appropriate instruction using the Spanish language in science, mathematics, social studies, and Spanish as a first language. English as a second language also was provided for students at all sites, as in previous years.

Pupils in Grades 3-12 were served, with the instructional format articulated to meet the demands of each instructional level. In the elementary school the program was operated in self-contained classrooms. In the junior high school, where pupils moved from class to class for each subject, students were in Bilingual Program classes for part of the day and received instruction with Anglo classmates for the remainder. In the senior high school students had individual rosters, permitting the selection of courses from among those offered by the ARRIBA component.

The ARRIBA program was operational in Ludlow, Waring and Southwark elementary schools, Stoddart-Fleisher and Penn Treaty junior high schools, and William Penn, Benjamin Franklin, Thomas Edison, and Kensington high schools. Southwark school was added to the program when classes from Waring elementary school were bussed there in order to improve integration.

Pupils Served

At the end of the 1972-1973 school year the three components of the Bilingual Program were serving a total of 1,834 students. Of those who answered the questionnaire, 1,110 said that they spoke mainly Spanish in their homes, 184 said that they spoke both Spanish and English, and 523 spoke mainly English. Seven hundred thirty-eight students were born in Puerto Rico, 925 were born on the United States mainland, and 31 were born in Spanish-speaking areas other than Puerto Rico.

There were 1,189 students in MODEL A, 571 speaking mainly Spanish, 88 speaking both Spanish and English, and 515 speaking mainly English. In MODEL B there were 170 students, 146 of whom spoke mainly Spanish at home and 24 of whom spoke both English and Spanish. Of the 475 students in the ARRIBA program, 393 spoke mainly Spanish at home, 72 spoke both Spanish and English, and 8 spoke mainly English.

School Disruptions

The program operation was disrupted by two teacher strikes, and the illness and death of one of the program supervisors.

Many class days during 1972-1973 were lost due to an extensive teachers' strike. Schools opened 17 days late on September 29, and were virtually closed for 38 days between January 8 and February 28. Sixteen days were made up by keeping schools open on holidays, during

regular vacation periods, and throughout the month of June. Because striking occurred at two different times during the year, instruction was segmented and students did not have an opportunity to really settle into a long period of school. The project director felt that two weeks of progress were lost for every week during which schools were closed. The evaluation was affected in that end-of-year testing was delayed as long as possible in an effort to allow a maximum of instructional time before tests were administered. The planned log study was omitted because the updating of micro-objectives could not be completed early in the year because teachers were striking, and because the time required to test the students with this instrument seemed excessive in view of instructional time already lost.

The program suffered a great loss in the illness and death of Ramona Rodriguez, who had supervised the ARRIBA teacher training and curriculum development. The temporary adjustments which were made in supervisory roles due to her illness could not compensate for the experience and skill which she brought to this role. Her loss will continue to be felt by all involved in the program.

Program Operational Cost

Table 1.1 shows the estimated expenditures from Title VII funds for operating the LET'S BE AMIGOS program (excluding evaluation and audit costs). These remained virtually the same as those occurring in the 1971-1972 school year ($165 per pupil) as rising salaries were offset by periods for which striking personnel were not paid and the payment of increased numbers of faculty teachers out of the School District's general operating budget. The figure represents a considerable reduction over the costs of two years ago (See Offenberg et al., 1973, p.iv) when the costs were about $340 per pupil.

Community Relations Activities

As most community-relations activities were reviewed in the 1973-1974 proposal, only an abstract is presented here. This abstract has been updated to include events through the end of the fiscal year.

During the 1972-1973 school year numerous of gatherings were attended by pupils, parents, and staff. The Bilingual Program initiated 10 holiday celebrations, six meetings involving bilingual teachers, 13 activities which shared information related to needs of the community and services being offered, and 10 meetings (some of which were for

TABLE 1.1 DIRECT EXPENDITURES FOR THE TITLE VII PROGRAM, "LET'S BE AMIGOS": 1972-1973, FROM FEDERAL FUNDS, EXCLUDING EVALUATION AND AUDIT EXPENDITURES

	Teaching Salaries*	Supervision, Administration, Curriculum Development**	Clerical Expenses**	Books, Teaching Materials**	Miscellaneous Expenses and Supplies**	Total	Cost per Pupil
MODEL A No. of Pupils= 1,189	$ 88,763	$102,283	$17,838	$20,154	$21.468	$250,506	$211
MODEL B No. of Pupils= 170	21,298	14,624	2,551	2,882	3069	44,424	261
ARRIBA No. of Pupils= 475	126,705	40,862	7,126	4,052	8,576	191,321	403
Total No. of Pupils= 1,834	$236,766	$157,769	$27,515	$31,088	$33,113	$486,251	$265

* These include salaries of teachers charged against Title VII; others are paid out of local funds.

** These expenses are pro-rated on the basis of the number of pupils in each program component.

planning) for parents of children in the Bilingual Program.

Ninety-eight activities of other community groups not directly related to the Bilingual Program received participation, including events to promote awareness, understanding, and cooperation in the community, cultural events, and meetings to give tribute to groups and individuals.

Curriculum Development

Objective 1.5 of the ARRIBA component stated that, where needed, curriculum materials would be developed for that component. Since the initiation of the program, curriculum materials have also been developed for the MODEL A and MODEL B programs, although no specific objective in the proposal mandated that this be done.

During the fiscal year ending in June 1973, curriculum writing involved 11 teachers working in the program and six teacher-specialists who were with the Philadelphia schools but not teaching in LET'S BE AMIGOS. The emphasis in curriculum development has changed from primarily development of new materials to adaptation of existing English-language materials where possible and preparation of new materials only when necessary. According to the coordinator of curriculum development, this change in emphasis was due, in part, to problems discussed ir Study 3, which focuses on evaluation of some original materials.

The following completed materials were adapted from School District of Philadelphia curricular materials and are available for tryout:

- LA VIDA EN LAS CIUDADES, PUEBLOS Y FINCAS DE PENSILVANIA. Social studies teacher's guide for the intermediate grades (4 through 6)

- COMO EL AMBIENTE FISICO AFECTA LA FORMA DE VIDA EN EL HEMISFERIO OCCIDENTAL. Social studies teacher's guide for the intermediate grades

- SCIENCE FOR THE INTERMEDIATE GRADES: ELECTRICITY AND

MAGNETISM (IN SPANISH). Teacher's guide

- SCIENCE FOR THE INTERMEDIATE GRADES: SOUND (IN SPANISH). Teacher's guide

- LA TIERRA EN EL ESPACIO (UNITS I,II, AND III). Teacher's guide for the intermediate grades

- SCIENCE FOR THE INTERMEDIATE GRADES: SUPPLEMENTARY MATERIAL FOR MAGNETISM, ELECTRICITY (and other units yet to be translated). Teacher's guide.

In addition to the adaptation of existing School District materials, one set of Spanish-as-a-first-language materials was adapted from other sources:

- LA LITERATURA PUERTORRIQUEÑA contains a short story, two plays, and a novelette, all adapted for use in the classroom.

The following original materials were released for distribution during 1972-1973:

- TU Y TU VIDA: LA SOLEDAD is an intermediate-grade material designed to give Spanish-speaking pupils information about dangerous drugs. It consists of a teacher's guide and a workbook.

- MIS PRIMEROS PASOS CON LAS VOCALES is a teacher's manual and two workbooks designed to develop reading readiness in Spanish (in the prekindergarten, kindergarten, and first grade). Development of these materials was begun several years ago, but the final version was prepared for release in 1972-1973.

- MIS PRIMEROS PASOS A LA LECTURA is a student workbook, designed for use with the primer levels of the Laidlaw Reading Series in Spanish. It is meant for kindergarten and first grade.

- PUERTO RICAN HISTORY AND CULTURE: A TEACHER'S RESOURCE

UNIT is an English-language set of materials to be used by teachers, both within and outside the program. It is designed for junior high school pupils.

As was the case in past years, a curriculum-evaluation checklist was developed for use in determining whether materials were being produced which met quality criteria established for the program. This year's checklist (shown in Table 1.2) contains some questions added to those used in the past, in order to improve control of the curriculum-development process (Questions 2, 3, 14, 15, 16 and 17). The ratings were made by the project's coordinator of curriculum development.

TABLE 1.2 COORDINATOR OF CURRICULUM DEVELOPMENT'S REVIEW OF 11 SETS OF MATERIALS COMPLETED AND RELEASED IN 1972-1973

Characteristic	No. of Units		
	Having Characteristic	Not Having Characteristic	Characteristic not Applicable
1. Appropriateness for intended grade level(s)	10	0	1
2. Intended grade level(s) indicated on cover	8	2	1
3. Subject matter area indicated (if not clear from title itself)	8	2	1
4. Appropriate for students' cultural background, interest level, and experiential field	9	1	1
5. Appropriate for students' previous knowledge in the subject matter	5	0	6

TABLE 1.2 (cont'd)

Characteristic	No. of Units Having Characteristic	No. of Units Not Having Characteristic	Characteristic not Applicable
6. Specific behavioral objectives clearly stated at beginning of each section or lesson	6	4	1
7. Sequential organization and structure	8	0	3
8. Testing procedures and materials included	3	7	1
9. Reasonable variety of learning activities	9	1	1
10. Provision for individual rate of learning	4	6.	1
11. Provision for enrichment or extended practice activities	7	3	1
12. Availability of audio-visual or other equipment needed to teach the unit	7	0	4
13. Aids and materials needed to teach unit: how teacher may make them or where available	7	0	4

TABLE 1.2 (cont'd)

---

14. Unit is

4 completely original

1 adapted from commercial material

6 translation of unit of School District of Philadelphia

15. Material is

6 teacher's guide

1 student workbook

4 both teacher's guide and student workbook

16. An additional translation of unit would be valuable

2 in English

0 in Spanish

9 translation already exists in needed language

17. Approval for copyrighted materials included in unit

0 needed

11 not needed

0 already obtained in writing

As can be seen in the table, the completed materials meet, for the most part, the demands of the curriculum-development process. The two areas in which most of the materials fail to meet them are Items 8 and 10: most materials did not include adequate testing procedures and most did not include provision for individual rates of learning. According to the curriculum-development coordinator, these deficiencies developed because the English materials used in producing the Spanish versions also lacked these elements. However, a lack of testing procedures was also characteristic of other project-developed materials in previous years.

Several other sets of materials were also being developed during 1972-1973, but were not yet completed or released for general distribution:

- ENGLISH: YOUR SECOND LANGUAGE, for ESL, in Grades 5-7
- EL RECIAN LLEGADO, GIGANTES BORINCANOS, UNA CARTA, BUSCANDO EMPLEO, SALIENDO DEL CINE, UN COMPROMISO FORMAL revision of Muckley units for SFL in Grades 10-12
- ROBERTO CLEMENTE--EL HOMBRE, EL ATLETA, EL HEROE, for in SFL, Grades 10-12
- LET'S READ, A PROGRAM OF READING READINESS IN THE ELEMENTARY GRADES for ESL in Grade 3
- CANCIONES INFANTILES for social studies in prekindergarten through Grade 2
- HABLEMOS, NINOS., Level D SFL in Grades 2-3
- TU Y TU VIDA PT. II for drug education and social studies in Grades 4, 5, 6
- MIS PRIMEROS PASOS CON LAS CONSONANTES Y LAS SILABAS for SFL in Grades K-2

Summer Institute

A Summer Teacher-Training Institute was held in conjunction with Temple University. The institute was designed to train native Spanish-speakers with two years of college. Upon completion of the program, participants were awarded emergency certification by the State of Pennsylvania. Twenty teachers were trained in the 1972 summer institute for Spanish-speaking professionals. All twenty completed the program and were employed in the various bilingual programs in the city.

Evaluators' Comment

The changes reported in the foregoing sections generally reflect evolution of the program and continued progress in fulfilling the plans set in the proposal documents. However, two events reported, the strikes and the death of Mrs. Rodriguez, were expected to have major impact on the performance of the student target groups. It is therefore important that the reader keep in mind that the 1972-1973 year was atypical--a fact that was likely to temporarily reduce the performance of the student participants.

STUDY 2

REORGANIZATION OF MODEL A AND MODEL B MICRO-OBJECTIVES IN LANGUAGE ARTS SCIENCE, SOCIAL STUDIES, AND MATHEMATICS INTO A CONTINUOUS SEQUENCE

Since the inception of the Bilingual Program LET'S BE AMIGOS, instruction has been guided by the Pupil Logs, a set of lists of micro-objectives used in MODEL A and MODEL B in both first-and second-language instruction in language arts, social studies, science, and mathematics. Each year new micro-objectives were added as the program cycled upward through the grades, and each grade level had its own specific list of items. Each year (see Offenberg 1971, 1972, 1973) students were tested with the list of micro-objectives specified for their grade, from pre-kindergarten to grade three. One problem with these logs was that they were not internally consistent, i.e., some objectives which were assigned to higher grade levels appeared to teachers to be easy, and others assigned to early grades appeared to be difficult. One aim of the present study is to put the objectives in order of difficulty.

During the 1971-1972 school year, it became apparent that the infusion of new students into the program at various grade levels reduced the fit of the grade-level-by-grade-level list of micro-objectives with the instructional program as it was being implemented.

These problems indicated a need for the adoption of a "continous primary" approach to program evaluation--an approach which permits pupils to work through their curriculum at a pace consistent with their past experiences and abilities.

When in the 1972-1973 school year prolonged teacher strikes forced a reduction in the amount of testing which could be carried out, application of the logs was temporarily dropped. Instead, a revision of the log was undertaken in anticipation of 1973-1974.

The purpose of this revision was to change the log instruments from a year-by-year format to a set of micro-objectives which increase monotonically in difficulty in each subject area.

To accomplish this task, teachers with expertise in each curricular area (language arts, science, social studies, and mathematics)

were asked to put objectives in what they considered to be the appropriate order.

Next year, the newly organized logs will be administered to pupils in various grade levels with various amounts of Bilingual Program experience. The revised log will then permit description of these diverse sets of student groups on a common scale, and thus enable the project staff to assess the impact of the program on the various groups.

Procedure

Evaluation

In order to reorganize the items in the logs, teachers who knew the subject areas represented by micro-objectives were asked to review them and order them from those which should be taught first to those which should be taught last. Supervisors designated a total of 16 teachers who were the most familiar with teaching the areas.

As shown in Table 2.1 there were between three and five teachers to rate the items in each subject area.

TABLE 2.1 NUMBER OF RATERS BY SUBJECT AREA

Subject Area	No. of Raters
Mathematics	4
Science	5
Language Arts	4
Social Studies	3
Total	16

In the original logs, some micro-objectives appeared more than once because the initial classification by language and grade level permitted repetition. When the repetitions were eliminated, there were 252 distinct concepts. As in the original classification of the micro-objectives, these concepts were categorized by content into four subject areas--mathematics, science, language arts, and social studies--and a deck of cards was prepared for each. The number of items in each of the four decks is shown in Table 2.2.

TABLE 2.2 NUMBER OF ITEMS BY SUBJECT AREA

Subject Area	No. of Items
Mathematics	74
Science	29
Language Arts	79
Social Studies	70
Total	252

Sufficient copies of the decks were made for the raters in the subject areas, and each deck was shuffled to remove any ordering of the micro-objectives. The teachers were asked to order the micro-objectives from those which should be taught first to those which should be taught last. They were told to do so in the way that would be fairest, regardless of the mother tongue of the grade level of the pupil to be tested. They worked independently, each teacher making the judgment without consultation with other teachers rating the same objectives.

Analysis

Analysis of the data was designed (a) to see if there was sufficient agreement among teacher-raters to determine a clear ordering which could be used in the program, and (b) to determine the best ordering once

this was established. The specific analysis for each of the four subject areas was as follows:

1. Each teacher's responses were ranked, assigning the number 1 to the micro-objective to be taught first, 2 to the one to be taught second, etc.

2. The Kendall coefficient of concordance W was calculated to assure that the ranks given by the teachers for each area were similar (See Segal, 1956, pp. 229-238).

3. The mean of the ranks assigned by the teachers to each item was calculated and the micro-objectives were reordered according to these means.

Results

As shown in Table 2.3, there was high agreement in the teachers' decisions in ordering micro-objectives from those to be taught first to those which should be taught later in the program.

TABLE 2.3 STRENGTH OF AGREEMENT AMONG RATERS

Subject	W	Chi-square*	df	p
Mathematics	.82	239.44	73	.001
Science	.85	119.00	28	.001
Language Arts	.87	288.84	83	.001
Social Studies	.90	186.30	69	.001

*Chi-square equivalents were calculated in order to show the level of significance of the coefficients.

To test the significance of W, chi-square equivalents were computed. The chi-square values in the table show that the agreement in the ranking of the items in each subject area was high enough to be statistically significant: none could be due to chance more than once in

1,000 times. Tables 2.4, 2.5, 2.6, and 2.7 show the micro-objectives for the four subject areas in order of their mean rank. The tables also show the original grade level at which each micro-objectives first appeared in the program of pupil work in the first language. Many of these objectives were introduced in later grade levels in the second language in the original logs.

TABLE 2.4. SEQUENCE OF MICRO-OBJECTIVES FOR LANGUAGE ARTS

	Micro-objective	Original Level*
1.	What is your name?	Prekindergarten
2.	Touch your eyes.	Prekindergarten
3.	Touch your nose.	Prekindergarten
4.	How old are you?	Prekindergarten
5.	Touch your feet.	Prekindergarten
6.	Touch your head.	Prekindergarten
7.	Identify mouth.	Kindergarten
9.	Identify ears.	Kindergarten
9.	Identify teeth.	Kindergarten
10.	Identify hands.	Kindergarten
11.	Identify face.	Special First Grade
12.	Identify fingers.	Kindergarten
13.	Identify arms.	Special First Grade
14.	Identify legs.	Special First Grade
15.	Where do you live? (Complete sentence)	Kindergarten
16.	When you look, you use your (eyes).	Prekindergarten
17.	When you run, you use your (legs)	Prekindergarten
18.	Identify red.	Prekindergarten
19.	When you listen, you use your (ears)	Prekindergarten
20.	Identify yellow.	Prekindergarten
21.	Sit down.	Kindergarten
22.	Identify blue.	Prekindergarten
23.	Stand up.	Kindergarten
24.	On what part of your body do you put your shoes?	Prekindergarten
25.	What do you do with your eyes?	Grade 2
26.	Run.	Prekindergarten
27.	Jump.	Prekindergarten
28.	What do you do with your legs?	Grade 2
29.	Identify green.	Prekindergarten
30.	What do you do with your ears?	Grade 2
31.	Identify black.	Kindergarten
32.	Hop.	Prekindergarten
33.	Identify white.	Kindergarten
34.	Skip.	Prekindergarten
35.	Identify pink.	Special First Grade
36.	Identify brown.	Kindergarten
37.	What color are your eyes?	Prekindergarten
38.	What is your address?	Prekindergarten
39.	Identify gold.	Special First Grade

TABLE 2.4 (cont'd)

	Micro-objective		Original Level*
40.	What color is your hair?		Prekindergarten
41.	Identify silver.		Special First Grade
42.	Show me_____.		Kindergarten
43.	Get me the_____.		Prekindergarten
44.	Give me the_____.		Prekindergarten
45.	Get the_____.		Kindergarten
46.	Who is he/she?		Prekindergarten
47.	Pass the_____.		Kindergarten
48.	Identify chair.		Special First Grade
49.	Identify T.V.		Special First Grade
50.	Where do you go to school?		Prekindergarten
51.	Identify bed.		Special First Grade
52.	Identify bedroom.		Special First Grade
53.	Identify kitchen.		Special First Grade
54.	Identify bathroom		Special First Grade
55.	Identify living room.		Special First Grade
56.	Identify dining room		Special First Grade
57.	Identify lamp.		Special First Grade
58.	Identify sofa.		Special First Grade
59.	Identify table.		Special First Grade
60.	Identify bathtub.		Special First Grade
61.	Identify sink.		Special First Grade
62.	Identify toilet		Special First Grade
63.	Identify refrigerator.		Special First Grade
64.	Identify stove.		Special First Grade
65.	Identify rug.		Special First Grade
66.	Identify capital letters in own name.		Kindergarten
67.	Print own name.		Kindergarten
68.	Identify toaster.		Special First Grade
69.	Identify mixer.		Special First Grade
70.	Identify vacuum cleaner.		Special First Grade
71.	Identify bureau.		Special First Grade
72.	Identify washbasin.		Special First Grade
73.	Identify buffet.		Special First Grade
74.	Identify china closet.		Special First Grade
75.	Word pairs beginning with same letter:	(mama-mono) (mama-libro)	Special First Grade
76.	Word pairs beginning with same letter:	(four-five) (four-six)	Special First Grade
77.	Given a word, list orally others with same initial sound.		Grade 2

TABLE 2.4 (cont'd)

Micro-objective	Original Level*
78. Orally, use word houses or people in a short sentence.	Grade 2
79. Given a word, list orally others with same ending sound.	Grade 2
80. Write sentences based on words written on the board.	Grade 2

*Some micro-objectives were reentered in several grade levels' logs. The earliest grade level in which the micro-objective appeared in the log is shown in this column.

TABLE 2.5 SEQUENCE OF MICRO-OBJECTIVES FOR MATHEMATICS

	Micro-objective	Original Level*
1.	Count from one to five.	Prekindergarten
2.	Count from 1 to 10.	Prekindergarten
3.	Which is heavier, the book or the pencil?	Prekindergarten
4.	Count from 1 to 12.	Kindergarten
5.	Recognize numerals from 1 to 10 shown out of sequence	Kindergarten
6.	How many items do you have? (Give 2 to 10 objects)	Prekindergarten
7.	Pick object which approximates a circle.	Prekindergarten
8.	Pick object which approximates a square.	Prekindergarten
9.	Pick object which approximates a triangle.	Prekindergarten
10.	Pick object which approximates a rectangle.	Prekindergarten
11.	Using objects, add one more. How many do you have now?	Prekindergarten
12.	Add two groups of objects (total less than 10)	Prekindergarten
13.	Match (children to chairs) (children to pencils)	Special First Grade
14.	Separate objects into similar pairs.	Kindergarten
15.	Count from 1 to 100.	Special First Grade
16.	9 + 4 =	Grade 2
17.	6 + 9 =	Grade 2
18.	5 + 8 =	Grade 2
19.	8 + 9 =	Grade 2
20.	15 - 6 =	Grade 2
21.	13 - 8 =	Grade 2
22.	17 - 9 =	Grade 2
23.	Identify nickel.	Special First Grade
24.	Identify dime.	Special First Grade
25.	Identify quarter.	Special First Grade
26.	What number comes before 20?	Grade 3
27.	Using a clock, indicate (one hour) (half hour)	Special First Grade
28.	How many days in a week.	Grade 2
29.	Identify (pint) when shown containers. (quart) when shown containers.	Grade 2
30.	2 x 7 =	Grade 2
31.	5 + 2 + 4 = (vertical)	Grade 3
32.	Tell time when shown a specific hour.	Special First Grade
33.	Count by 2 up to 10.	Grade 2
34.	Count by 5 up to 30.	Grade 2
35.	3 x 4 =	Grade 2

TABLE 2.5 (cont'd)

	Micro-objective	Original Level*
36.	Count by 10 up to 100.	Grade 2
37.	5 x 3 =	Grade 2
38.	How many months in a year?	Grade 2
39.	14 ÷ 7 =	Grade 2
40.	Divide a circle into halves.	Grade 2
41.	How many halves in a whole?	Grade 2
42.	12 ÷ 4 =	Grade 2
43.	Which one shows ¼ of a circle?	Grade 3
44.	Divide a square into fourths.	Grade 2
45.	15 ÷ 5 =	Grade 2
46.	How much of the circle is dark?	Grade 3
47.	18 - 2 = (vertical)	Grade 3
48.	½ of 6 =	Grade 2
49.	½ of 8 =	Grade 2
50.	How many minutes in an hour?	Grade 2
51.	How many quarters in one dollar?	Grade 3
52.	2 x 4 = (vertical)	Grade 3
53.	31 - 3 = (vertical)	Grade 3
54.	Identify tablespoon.	Grade 2
55.	How many 10's in 32?	Grade 3
56.	2 12 =	Grade 3
57.	Use of decimal notation of money: $.10	Grade 2
58.	Use of decimal notation of money: $.25	Grade 2
59.	Count by 5's up to 50.	Grade 3
60.	12 x 2 = (vertical)	Grade 3
61.	Read the following aloud: $2.60	Grade 3
62.	22 + 36 + 25 = (vertical)	Grade 3
63.	How many 1's, 10's, and 100's in 125?	Grade 3
64.	9 x 8 - (vertical)	Grade 3
65.	3 64 =	Grade 3
66.	Read the following aloud: 253	Grade 3
67.	Count backward from 50 to 10 by 10's.	Grade 3
68.	Read the following aloud: 845	Grade 3
69.	What number is 2,635?	Grade 3
70.	Add: 2.43 + 1.51 (vertical)	Grade 3
71.	What comes before 425?	Grade 3
72.	(2 x 10) + (5 x 1) - (accept (20+5) or 25	Grade 3
73.	Read the following Roman number: VI	Grade 3
74.	Read the following Roman number: IX	Grade 3

*Some micro-objectives were reentered in several grade levels' logs. The earliest grade level in which the micro-objective appeared in the log is shown in this column.

TABLE 2.6 SEQUENCE OF MICRO-OBJECTIVES FOR SCIENCE

	Micro-objective	Original Level*
1.	Identify dog from picture.	Prekindergarten
2.	Identify cat from picture.	Prekindergarten
3.	Identify bird from picture.	Prekindergarten
4.	Identify apples from picture.	Prekindergarten
5.	Identify horse from picture.	Prekindergarten
6.	Identify cow from picture.	Prekindergarten
7.	Identify bananas from picture.	Prekindergarten
8.	Identify oranges from picture.	Prekindergarten
9.	What do we wear when it's hot?	Prekindergarten
10.	What do we wear when it's cold?	Prekindergarten
11.	Identify potatoes from picture.	Prekindergarten
12.	Identify sheep from picture.	Prekindergarten
13.	Place correct weather symbol on chart.	Prekindergarten
14.	Identify roses from picture.	Prekindergarten
15.	Identify lion in a sentence.	Kindergarten
16.	Identify tiger in a sentence.	Kindergarten
17.	Identify bear.	Special First Grade
18.	Identify duck.	Special First Grade
19.	Identify turtle.	Special First Grade
20.	Identify chicken in a sentence.	Kindergarten
21.	Identify gorilla.	Special First Grade
22.	Identify goat.	Special First Grade
23.	Identify frog.	Special First Grade
24.	Identify cherries.	Special First Grade
25.	Identify spinach	Special First Grade
26.	Identify daisies.	Special First Grade
27.	What dissolves in H_2O?	Grade 2
28.	What does not dissolve in H_2O?	Grade 2
29.	How does the weather change?	Grade 2

*Some micro-objectives were reentered in several grade levels' logs. The earliest grade level in which the micro-objective appeared in the log is shown in this column.

TABLE 2.7 SEQUENCE OF MICRO-OBJECTIVES FOR SOCIAL STUDIES

	Micro-objective	Original Level*
1.	Identify policeman from picture.	Prekindergarten
2.	Identify police car.	Kindergarten
3.	Match the pictures of policeman-police car	Special First Grade
4.	Why is the police car important to the policeman?	Grade 1
5.	Identify doctor from picture.	Prekindergarten
6.	Identify fireman.	Prekindergarten
7.	Match the pictures of fireman-fire engine.	Special First Grade
8.	Identify car from picture.	Prekindergarten
9.	Identify nurse from picture.	Prekindergarten
10.	Identify fire engine.	Kindergarten
11.	What is the ambulance used for?	Grade 2
12.	Who do we call when a fire starts?	Grade 2
13.	How does the fireman help us?	Grade 2
14.	What is the fire engine used for?	Grade 2
15.	Why is the fire engine important to the fireman?	Grade 1
16.	Identify milkman from picture.	Prekindergarten
17.	What school are you in?	Special First Grade
18.	Identify milk truck.	Kindergarten
19.	What is your teacher's name?	Special First Grade
20.	Match the pictures of milkman-milk truck.	Special First Grade
21.	What grade are you in?	Special First Grade
22.	Why is the milk truck important to the milkman?	Grade 1
23.	What days do we come to school?	Kindergarten
24.	Identify mailman.	Kindergarten
25.	What is your mother's name?	Special First Grade
26.	What is your father's name?	Special First Grade
27.	Identify newspaper boy.	Special First Grade
28.	Identify garbage collector.	Special First Grade
29.	How many sisters and brothers do you have?	Kindergarten
30.	Identify and talk about family members from picture.	Special First Grade
31.	What is your sister's/brother's name?	Special First Grade
32.	How does the trashman help us?	Grade 2
33.	Identify bus from picture.	Prekindergarten
34.	Where do we go when we want to buy stamps?	Grade 2
35.	Identify trolley from picture.	Prekindergarten
36.	Where do we go when we want books?	Grade 2

TABLE 2.7 (cont'd)

	Micro-objective	Original Level*
37.	Describe use of traffic light.	Grade 2
38.	Describe use of stop sign.	Grade 2
39.	Why don't we go to the shore in the winter?	Grade 3
40.	Kinds of weather in winter but not in summer in Philadelphia	Grade 3
41.	Difference between winter and summer clothes in Philadelphia.	Grade 3
42.	What is the difference between party clothes and school clothes?	Grade 3
43.	What do plants need to grow?	Grade 3
44.	Name some foods made from milk.	Grade 3
45.	Name some foods that grow underground	Grade 3
46.	What animals give us wool	Grade 3
47.	Why do some foods come in cans and some frozen?	Grade 3
48.	How is manufactured clothing brought to our stores stores?	Grade 3
49.	Why are most trees green in Puerto Rico but lose their leaves in Philadelphia?	Grade 3
50.	What is a farm?	Grade 3
51.	Ho do farms help cities and towns?	Grade 4
52.	What machines do people have in their homes?	Grade 3
53.	What is a city?	Grade 4
54.	Name some things we use that are found in the earth.	Grade 3
55.	Name some things made of iron.	Grade 3
56.	What is an island?	Grade 4
57.	Who was the founder of Philadelphia?	Grade 4
58.	What is a town?	Grade 4
59.	Name some things with motors in them.	Grade 3
60.	What is the capital of Pennsylvania?	Grade 4
61.	Name one bridge connecting Philadelphia and New Jersey.	Grade 4
62.	How do you get to Philadelphia from New York?	Grade 4
63.	Who discovered Puerto Rico?	Grade 4
64.	What are the different ways of getting the news to people?	Grade 4
65.	What is the capital of Puerto Rico?	Grade 4

TABLE 2.7 (cont'd)

	Micro-objective	Original Level*
66.	Name two large cities in Puerto Rico.	Grade 4
67.	What is the typical dish of Puerto Rico?	Grade 4
68.	What is a popular musical instrument in Puerto Rico?	Grade 4
69.	Mention some typical products of Puerto Rico.	Grade 4
70.	What is the Three Kings' Day?	Grade 4

*Some micro-objectives were reentered in several grade levels' logs. The earliest grade level in which the micro-objective appeared in the log is shown in this column.

Evaluators' Comment

The high level of agreement among raters shows that experience in the program has led to a clear picture of the order in which material should be taught. It is interesting to note, however, that in some subject areas there is great disagreement with the original statement of goals as reflected in the year-by-year logs.

Language Arts

As bilingual education focuses on language performance, and as there is heavy emphasis on language development in the early grades of school, it is not surprising that many of the micro-objectives were in this category. It is more surprising that teachers were in sharp disagreement with the original logs in the order in which the concepts in this area were to be taught. Although there was a tendency for kindergarten and prekindergarten micro-objectives to appear more often in the first half of the instructional sequence and for first- and second-grade micro-objectives to appear later on, there was a substantial amount of mixing across the list. This suggested that the original ordering of micro-objectives in this area was not an appropriate guide for instruction of the pupils in the Model School components.

Mathematics and Science

In contrast to language arts, the mathematics and science teachers' ordering of micro-objectives was fairly similar to that of the original logs.

In science, supervisors' review of the total number of distinct objectives showed that this area was probably in need of enrichment. The 29 micro-objectives dealt primarily with identification and definitions, with only a few dealing with basic concepts which underlie the science curriculum used in the schools. It appears that additional attention is required in the science area.

Social Studies

Examination of the social studies micro-objective list showed that there was general agreement between the original logs and the teachers' ratings regarding which material should be introduced early in the instructional sequence and which material should be introduced later. However, within the two halves of the sequence, there was considerable mixing of grade levels.

Scope of the Objectives

In each area except social studies, there appeared to be little material which was designated in the original logs as work to be done in the fourth grade in the mother tongue. It is not clear to the evaluators whether this was deliberate or attention needed to be given to adding more complex skills to the language arts, science, and mathematics areas.

Conclusions

This study suggested that teachers want reorganization of the original curriculum content's sequence, and that some areas may require addition of more complex items.

The high level of agreement among raters showed that there should be little difficulty in adopting a "continuous primary" approach to describing pupil performance in the curricular areas.

STUDY 3

EVALUATION OF THE JUNIOR HIGH SCHOOL SOCIAL STUDIES AND SCIENCE MATERIALS THROUGH CRITERION TESTING OF STUDENT PERFORMANCE

A major aspect of the ARRIBA component is development and tryout of curriculum materials. This study deals with tryout of two sets of materials for seventh-grade courses in social studies and science. Previous study of materials developed by program staff of the ARRIBA component has been informal. In the third operational year of the program (1971-1972) the book Historia de Puerto Rico, assessed in this paper, was used by one teacher at the upper elementary level and one at the junior high school level. The elementary teacher felt that the materials were too difficult for her class. At the junior high school level it appeared that these materials needed supplementation and enrichment from other sources, but otherwise they were satisfactory.

During 1971-1972 one of the two sets of materials in the science package tested in this study was also tried: the set of six booklets on simple machines (Máquinas Simples). Teachers felt that these booklets were most suitable for ninth-grade pupils although the content was originally designed for seventh grade. Teachers felt that concrete demonstrations were needed with these materials, but there was no provision for them and no laboratory space. A member of the research staff familiar with the concepts presented in Máquinas Simples felt that some basic distinctions (e.g., between weight and mass) were unclear.

During 1971-1972, a second set of science materials containing three units--Our Solar System, Our Atmosphere, and Our Weather--became available but was not formally tried. These materials, which together are called Ciencia en Español, were given a systematic trial during 1972-1973 as part of this study of seventh-grade science materials.

In contrast to the earlier assessments of curriculum materials, which came from the teacher's view, the 1972-1973 evaluation focuses on performance of pupils in the seventh grade on criterion-referent tests designed to determine whether pupils had the knowledges and concepts which the materials were expected to teach.

As was stated in the 1971-1972 Continuation Proposal (page 111), the tests were designed to have a 60% minimum passing score, and a 75% correct mean score after correction for guessing. The latter score served as the criterion for pupil performance.

During the development of these units, a debate occurred within the program regarding the effect of reading skill on tests which might be prepared to test pupil performance in science and social studies. The design which follows deals with the question of impact of reading skill on performance on tests, as well as assessment of the students' success in meeting the objective.

Procedures

Program

During the summer of 1972 outlines of program content for seventh-grade science and junior high schcol history of Puerto Rico were prepared. In October, when schools opened after the first strike, the two teachers who were responsible for seventh-grade science and the two teachers responsible for seventh-grade social studies were contacted by the Coordinator of Curriculum Development and the Project Director. Both sets of teachers agreed to teach the content of the respective materials which had been developed for the project:

- Máquinas Simples (Simple Machines) Introduction and 5 Booklets: Plano Inclinado, La Rueda y Polea, Otros Modificaciones de la Rueda, Las Palancas, Trabajo y Fuerza

- Ciencia en Espanol: Nuestro Sistema Solar, Nuestra Atmosfera, El Tiempo (Science in Spanish: Our Solar System, Our Atmosphere, and The Weather)

- Historia de Puerto Rico (History of Puerto Rico).

It was originally felt that pupil work with the science materials would be completed by March 1973 and the material in social studies would be completed by February 1973. When a second strike occurred in January and February, completions of the sets of materials were delayed until May. As a result the testing took place in early June.

It was expected that the project-developed science materials would constitute the main structure of the science program, with the teacher bringing in supplementary demonstrations. Historia de Puerto Rico was designed to be taught along with related historical concepts, which teachers introduced along with the project-developed curriculum. Teaching was to take place not only with the written materials, but also with verbal presentation and review to assure that pupils with reading difficulties could master the content.

Evaluation

The evaluation procedure was designed to test pupil performance on the materials in order to see if they met the expectations of curriculum planners. It also attempted to show whether students could read the test content.

Method. During the summer of 1972, a member of the evaluation staff prepared a list of objective questions in Spanish covering each concept in the science package and Historia de Puerto Rico. The questions were designed to be a complete review of every fact and point in each set of instructional materials. In other words, if a pupil could answer all of the questions on these lists, he knew every fact or concept in the material.

Teachers were provided with sets of the instructional materials and a copy of the appropriate questions. They were told that they were to teach the materials (during the time periods specified in the "Program" section) to their seventh-grade classes. After completion of the materials, pupils were to be given a test based on the question list; the form of the question might be changed, but the content would not be altered.

In June 1973, the date adopted after the strike, a member of the research staff visited the classes at a prearranged time and administered the criterion tests to all students present on the day of the testing.

Instrument. One test was developed for each subject area, history of Puerto Rico and science. A member of the research staff rewrote each question in the original list so that all questions were in either true/false or multiple-choice format. The research staff and the coordinator of curriculum development then made the final item selection. The exam was prepared in the form of three or four subtests.

The science exam consisted of 20 multiple-choice items and two groups of ten true/false items. The history exam consisted of two groups of ten multiple-choice items and two groups of ten true/false items. Both tests are shown appended to this study (Appendix 3.1).

Where the items were divided into two groups (i.e., the science true/false items and the history multiple-choice and true/false items), this was done to permit assessment of the effect of students' reading skills on test performance. Items judged to be of similar difficulty were assigned to either of two groups--in one group to be read aloud and in the other to be completed by the students' own silent reading of the questions.

Subjects. This testing was conducted late in June, beyond the normal closing date of school, and no makeup date could be scheduled. Many pupils were not attending school presumably because they were visiting their relatives in Puerto Rico or had left the city for some other reason. Absence for these reasons was added to normal absenteeism. As a result, 55% of the students on roll were tested in science and 60% of the students were tested in social studies.

Administration. The tests were administered in the regular classroom meeting of the students. A member of the research staff was present to monitor. First students were given those parts of the test which the pupils both saw and heard. With these tests in front of the students, the regular classroom teacher read the questions aloud and the students marked their answers. Then the subtests which the students were to read without help were distributed, and pupils worked the items silently. The pattern of administration in the two junior high schools was as shown in Table 3.1. This pattern is counterbalanced so that the effect of reading can be teased out for all but the multiple-choice items of the science test.

TABLE 3.1. PATTERN OF CRITERION SUBTEST ADMINISTRATION

Subtest	Stoddart-Fleisher	Penn Treaty
Social Studies		
Multiple-Choice 1	Read Aloud	Silent Only
Multiple-Choice 2	Silent Only	Read Aloud
True/False 1	Read Aloud	Silent Only
True/False 2	Silent Only	Read Aloud
Science		
Multiple-Choice	Silent Only	Silent Only
True/False 1	Read Aloud	Silent Only
True/False 2	Silent Only	Read Aloud

Analysis. Subtest and total scores were calculated, taking into account the effect of guessing. A tabulation of difficulty of each item was prepared which showed the percentage of pupils succeeding on each item. Where differences in means of equivalent parts were found which suggested that the oral presentation of questions had an effect, a t test was computed to show whether the differences were greater than chance.

Results

Observation

Classroom observation of teachers by the evaluation staff at the midpoint of the school year (five observations in all) showed that the materials were in use, but that one of the history teachers had problems of class control and student disinterest. Prior to testing, the science teacher at Penn Treaty stated that despite the emphasis on conforming to the printed schedule, he did not teach the material concerning pulleys and inclined planes. He expected that most students would therefore be unable to do Items 2 and 5 on the multiple-choice and Item 9 on each of the two groups of true/false items.

Testing of Science Materials

Table 3.2 shows the number of items on each science subtest, the score which could be obtained by guessing, the minimum number needed to "pass" the test--that is, to score 60% correct beyond the guessing level--and the number of items to score 75% beyond the guessing level. Table 3.3 shows the actual performance of pupils on each subtest. As can be seen in the table, only one of the 31 pupils tested had an overall score above the

level for passing the test, and all the group means were well below those anticipated in the original objective. Means for seventh graders on the multiple-choice section showed a two-item difference between the schools, probably due to failure to teach material related to two items on the test. Scores on the true false sections of the test were all approximately one item over the guessing level, suggesting that few pupils could really handle the content of these items.

Comparison of the group means for the parts which were read silently by some students and read aloud to others showed differences of 0.6 item or less in favor of the group which both heard and read the items, a difference which seemed too small to be educationally meaningful or to warrant further analysis.

TABLE 3.2. CHARACTERISTICS OF THE SCIENCE CRITERION TEST

Subtest	No. of Items	Guessing Score	Passing Score** (60%)	Objective Criterion Score** (75%)
Multiple-Choice	19*	6.3	14	16
True/False 1	10	5.0	8	9
True/False 2	10	5.0	8	9
Total	39	16.3	30	34

*Item 4 of the original 20 items was eliminated because the illustrator neglected to include the correct answer in the choices.

**Rounded to the nearest whole item.

TABLE 3.3. RESULTS OF SCIENCE TESTING IN SEVENTH GRADE

Subtest	Penn Treaty (N=19)	Stoddart-Fleisher (N=12)	Total (N=31)
Part I (Multiple-Choice)			
Mean Score	11.6	9.9	10.9
Standard Deviation	3.4	3.9	3.6
No. Passing Subtest	4(21%)	1(8%)	5(16%)
Part II (True/False 1)			
Mean Score	5.8	6.1*	5.9
Standard Deviation	1.9	1.4	1.7
No. Passing Subtest	3(16%)	2(16%)	5(16%)
Part III (True/False 2)			
Mean Score	6.1*	5.8	6.0
Standard Deviation	1.7	2.3	1.9
No. Passing Subtest	3(16%)	3(25%)	6(19%)
Total			
Mean Score	23.5	21.8	22.9
Standard Deviation	5.7	6.1	5.9
No. Passing Test	1(5%)	0(0%)	1(3%)

*Test parts which were read to pupils.

Taken together, the results of the testing of students on the science curriculum package showed that performance was far below levels expected by the curriculum development staff, and only very slightly above that which would be obtained by guessing on some test parts.

Testing of Social Studies Materials

Table 3.4 shows the characteristics of the social studies criterion test. As can be seen on the table, it consisted of two groups of

TABLE 3.4. CHARACTERISTICS OF THE SOCIAL STUDIES CRITERION TEST

Subtest	No. or Items	Guessing Score	Passing Score (60%)	Objective Score (75%)
True/False 1	10	5	8	9
True/False 2	10	5	8	9
Multiple-Choice 1	10	3.3	7	8
Multiple-Choice 2	10	3.3	7	8
Total	40	16.6	30	34

true/false items and two groups of multiple-choice items. "True/false" Part 1 and "Multiple-choice" Part 1 were read aloud to the Penn Treaty students as they worked the test; the students worked "True/false" Part 2 and "Multiple-choice" Part 2 silently. At Stoddart-Fleisher, the manner of presentation of parts was reversed--"True/false" Part 2 and "Multiple-choice" Part 2 were read aloud.

Results of the testing are shown in Table 3.5. As can be seen in the table, the social studies test also proved very difficult for the pupils in the program, with only four students of the 33 who completed the test obtaining a passing overall score.

As there seemed to be some overall difference in subtest means for the two schools, the difference between the total of Parts 1 and 3 and the total of Parts 2 and 4 was computed for each student. If there was a difference in performance brought about by reading items aloud, at Penn Treaty the differences should be larger than at Stoddart-Fleisher (when sign is taken into account). A t test between these two sets of differences was computed. It showed that there was no significant difference ($t=0.53$, $df=31$). Therefore, it appeared that ability to read the items was not critical in determining pupil performance.

TABLE 3.5. RESULTS OF SOCIAL STUDIES TESTING IN SEVENTH GRADE

Subtest	Penn Treaty (N=20)	Stoddart-Fleisher (N=13)	Total (N=33)
True/False 1			
Mean Score	5.1	5.6	5.5
Standard Deviation	1.5	1.3	1.1
No. Passing Subtest	1(5%)	0(0%)	1(3%)
True/False 2			
Mean Score	5.9	6.5	6.1
Standard Deviation	1.9	0.9	1.6
No. Passing Subtest	4(25%)	0(0%)	4(12%)
Multiple-Choice 1			
Mean Score	7.0	5.4	6.3
Standard Deviation	1.9	2.0	2.0
No. Passing Subtest	11(55%)	1(8%)	12(36%)
Multiple-Choice 2			
Mean Score	6.8	6.2	6.5
Standard Deviation	2.0	2.6	2.2
No. Passing Subtest	12(60%)	5(38%)	17(51%)
Total			
Mean Score	24.7	23.7	24.3
Standard Deviation	4.2	5.9	4.9
No. Passing Test	2(10%)	2(15%)	4(12%)

Examination of subtest performance showed that on the true/false items, three of the four school subtest means were within one item of the guessing level. Pupils fared somewhat better with the multiple-choice items, where the means were at or close to the minimum passing scores. Even here, however, performance was below the 75% average score set by the criterion.

Evaluators' Comment

The results of these two studies of curriculum developed by the project suggest that there are serious problems in the chain which begins with preparation of these materials and ends with pupil performance. Given the clarity with which the project personnel specified early in the year exactly what was to be taught and tested, and the manner in which the tests were constructed, it seemed unlikely that the low performance could be attributed to the teachers' misunderstanding of what would be tested. That leaves two possible problems: either (a) the materials were unsatisfactory for teaching the desired concepts to children in the program, or (b) the teachers were ineffective in teaching the content of the materials to their classes.

If these materials are to be used again with students in the program, it is important that there be regular observation of the teaching procedures, regular checks by project management to see that the students are in fact being taught concepts of the materials in a satisfactory manner, and practical assistance for those teachers who need help.

Before they are used again, it would be important to examine the materials themselves in order to determine whether they provide clear explanations of the facts and concepts involved. To this end, Appendix 3.1, which shows the sample tests, also indicates the page in the materials most in need of rewriting.

APPENDIX 3.1

SAMPLES OF TESTS USED TO ASSESS THE "SCIENCE" AND "HISTORY OF PUERTO RICO" MATERIALS

Included are (a) the percentage of pupils getting each item correct, and (b) the source of each item in the texts.

EVALUACION DE LA UNIDAD "HISTORIA DE PUERTO RICO"

Nombre _____ Fecha _____

Escuela _____ Grado _____

Selección Múltiple

Escoge la mejor contestación y escribe la letra de ésta en el espacio a la izquierda:

Percent Correct

93 1. Las letras "F" e "I" en el escudo de Puerto Rico representan

A. Francia e Italia.
B. franjas e islas.
C. Fernando e Isabel.

p. 4

55 2. La aduana es una agencia del gobierno que controla

A. la entrada y salida de personas y productos.
B. el ejército.
C. la rivalidad entre países fuertes.

p. 12

33 3. Dice que no deben formar colonias en Latinoamérica otros países la

A. Doctrina de Independencia.
B. Doctrina de Monroe.
C. Doctrina de Truman.

p. 21

88 4. Palabras tales como bohío y yuca provienen de los

A. españoles.
B. franceses.
C. indios.

p. 33

64 5. Además de escribir el poema "Canto a Puerto Rico," José Gautier Benítez escribió:

A. el ensayo "La Alianza para el Progreso."
B. una colección de supersticiones.
C. artículos sobre la esclavitud.

p. 47

EVALUACION DE LA UNIDAD "HISTORIA DE PUERTO RICO"

Nombre _____ Fecha _____

Escuela _____ Grado _____

Selección Múltiple

72 6. Se considera que Puerto Rico ha sido oficialmente bilingüe después de

A. La Revolución Francesa.
B. La Guerra Hispanoamericana.
C. La Segunda Guerra Mundial.

P. 56

58 7. El gobierno de Puerto Rico tiene tres ramas:

A. la fraternidad, la igualdad, la libertad.
B. la isla, la ciudad, la aldea.
C. la leglislativa, la ejecutiva, la judicial.

P. 8

42 8. Los puertorriqueños son ciudadanos de Estados Unidos de Norteamérica desde el año

A. 1717.
B. 1817.
C. 1917.

P. 86

79 9. El baile clásico de la cultura puertorriqueña es

A. la danza.
B. la jota aragonesa.
C. el vals.

P. 83

45 10. El Dr. Agustín Stahl, nacido y criado en Alemania, combatió

A. el vómito negro.
B. el sarampión alemán.
C. la enfermedad de la caña de azúcar.

P. 76

EVALUACION DE LA UNIDAD "HISTORIA DE PUERTO RICO"

Nombre _____ Fecha _____

Escuela _____ Grado _____

I. Cierto o falso

Escribe en el espacio en blanco la letra "C" si la oración es cierta o "F" si es falsa:

Percent correct

33 1. No se sabe quién fue el autor de "La Borinqueña." p. 9

61 2. La palabra jíbaro proviene del nombre de una tribu africana. p. 36

70 3. La religión nunca ha tenido gran influencia en las fiestas y diversiones puertorriqueñas. p. 41

45 4. Una de las diversiones durante una típica fiesta patronal en Puerto Rico es el palo ensebao. p. 46

70 5. Se ha perdido casi toda la obra renombrada de José Campeche, incluso "El Velorio." p. 52

39 6. Un ejemplo de una copla del baile del seis chorreao es "La Noche de San Juan" de Lope de Vega. p. 53

85 7. Desde pequeño don Luis Muñoz Rivera se interesó por el bienestar de la gente de Puerto Rico y del mundo. p. 87

73 8. Para el siglo XIX muchos puertorriqueños querían acabar con el coloniaje que tenía España en Puerto Rico. p. 27

88 9. Luis Muñnoz fue un famoso santero. p. 58

58 10. Los pueblos de Puerto Rico se fundaron mayormente alrededor de la iglesia. p. 14

 EVALUACION DE LA UNIDAD "HISTORIA DE PUERTO RICO"

Nombre _____ Fecha _____

Escuela _____ Grado _____

I. Cierto o falso

Escribe en el espacio en blanco la letra "C" si la oración es cierta o "F" si es falsa:

Percent correct

36 1. Un resultado de la Guerra Hispanoamericana fue la independencia de Puerto Rico. p. 30

24 2. Los jíbaros de hoy en día siguen usando una forma de costumbrismo criollo para expresarse. p. 36

45 3. En el siglo XVI la importación a la isla de muchas imprentas tuvo como resultado la publicación de mucha literatura puertorriqueña. p. 47

76 4. Se considera a Juan Morell Campos como el sobresaliente genio musical de Puerto Rico. p. 50

58 5. Aunque José Campeche es el primer pintor puertorriqueño, también se le conoce como aficionado de la música. p. 63

33 6. Durante las fiestas de San Juan se bailaba el seis, un baile de seis personas. p. 55

39 7. A Eugenio de Hostos se le conoce más como político que como educador. p. 86

64 8. Se le considera a José Gautier Benítez como el poeta puertorriqueño más grande del siglo XIX y de todos los tiempos. p. 47

79 9. Cuba es una posesión de Estados Unidos. p. 30

45 10. Actualmente Puerto Rico es un país independiente. p. 30

EVALUACION DE LA UNIDAD "HISTORIA DE PUERTO RICO"

Nombre _____ Fecha _____

Escuela _____ Grado _____

Selección Múltiple

Escoge la mejor contestación y escribe la letra de ésta en el espacio a la izquierda:

Percent Correct

85 1. La bandera de Puerto Rico fue adoptada por

A. un grupo de patriotas.
B. Betsy Ross.
C. los españoles.

p. 8

82 2. El nombre del famoso pirata puertorriqueño tan amado y respetado por la gente es

A. Fajardo.
B. Arizmendi.
C. Cofresí.

p. 21

36 3. Abolieron la esclavitud en Puerto Rico

A. las Cortes Españolas.
B. Roberto Cofresí y David Porter.
C. Fernando e Isabel.

p. 26

85 4. Un jíbaro es

A. un estilo de arte colonial.
B. un campesino puertorriqueño.
C. un estilo de música puertorriqueña del siglo XIX.

p. 36

64 5. El santo protector de la isla de Puerto Rico es

A. Juan Ponce de León.
B. San Juan Bautista.
C. Santiago de Compostela.

p. 55

EVALUACION DE LA UNIDAD "HISTORIA DE PUERTO RICO"

Nombre _____ Fecha _____

Escuela _____ Grado _____

Selección Múltiple

Escoge la mejor contestación y escribe la letra de ésta en el espacio a la izquierda:

76 6. El ilustre filósofo y abogado Segundo Ruiz Belvis luchó a favor de

A. la Operación Manos a la Obra.
B. la abolición de la esclavitud.
C. las Cortes Españolas.

p. 26

70 7. El nombre oficial de Puerto Rico es

A. la República Demócrata de Puerto Rico.
B. el Estado Libre de Puerto Rico.
C. el Reino Libre de Puerto Rico.

p. 8

30 8. Durante los tres siglos de colonización en Puerto Rico la mayoría de los habitantes de la isla vivían en

A. las cuatro ciudades más grandes.
B. Culebra, Mona y Vieques.
C. la zona rural.

p. 17

58 9. Ramón Emeterio Betances, el Lincoln de Puerto Rico, organizó

A. la rebelión del Grito de Lares.
B. el Seminario Conciliar.
C. la Sociedad Económica de Amigos del País.

p. 73

51 10. Se le conoce como un gran orador que entusiasmaba a las multitudes con sus discursos a

A. Juan Ponce de León.
B. José de Diego.
C. Mariana Bracetti.

p. 68

EVALUACION DE LAS UNIDADES "MAQUINAS SIMPLES"

Nombre _____ Fecha _____

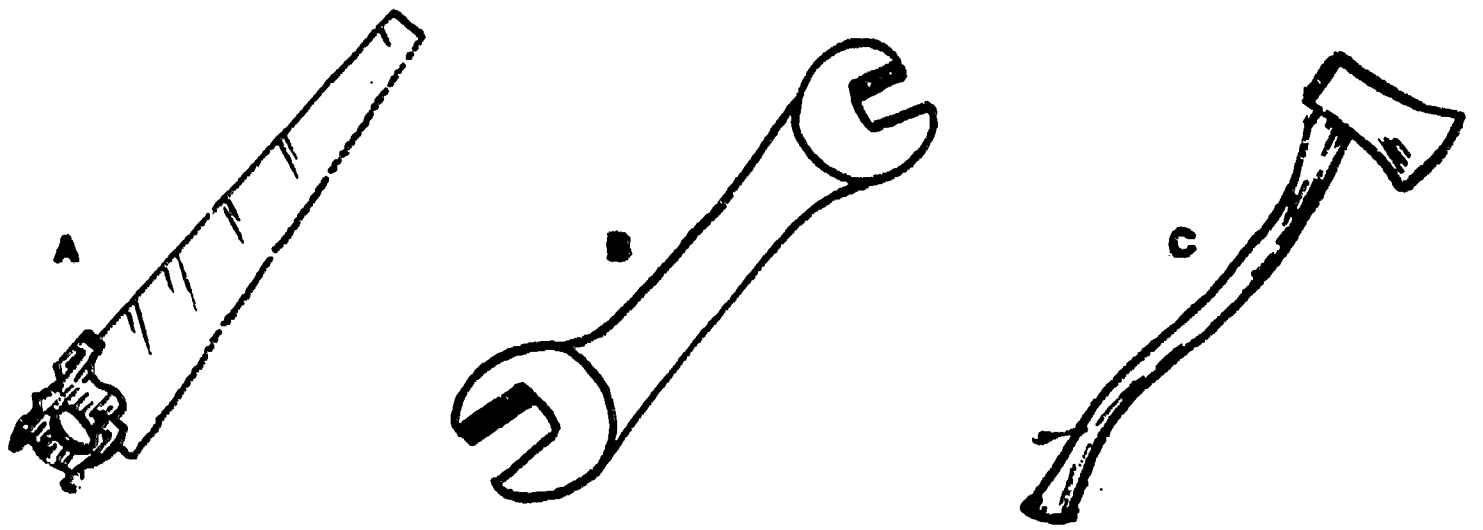
Escuela _____ Grado _____

Selección Múltiple

Escribe un círculo alrededor de la letra correcta:

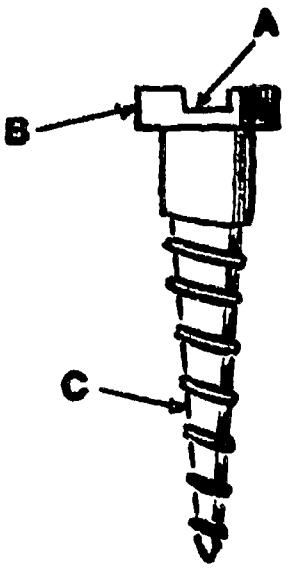
Percent Correct

(100) 1. Para desaflojar una tuerca debe usarse:



Mag. Simp: Intro.-p. 1

(13) 2. ¿Qué parte del tornillo es un plano inclinado?



Mag. Simp: Plano Inclinado,-p 6

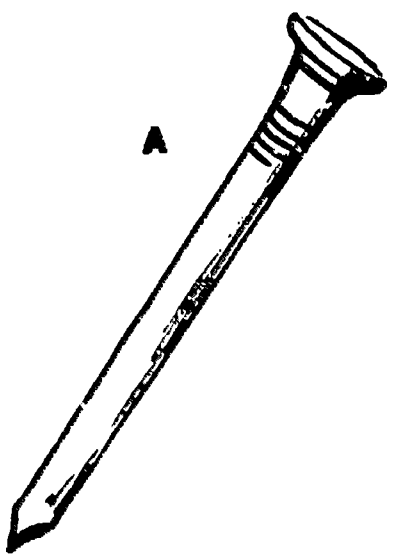
EVALUACION DE LAS UNIDADES "MAQUINAS SIMPLES"

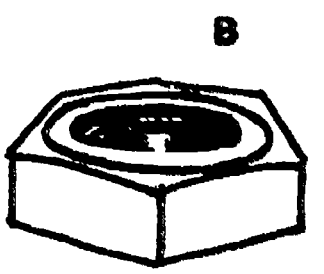
Nombre_____ Fecha_____

Escuela_____ Grado_____

Seleccion Múltiple (ctda)

(19) 3. El destornillador se usa para atornillar:

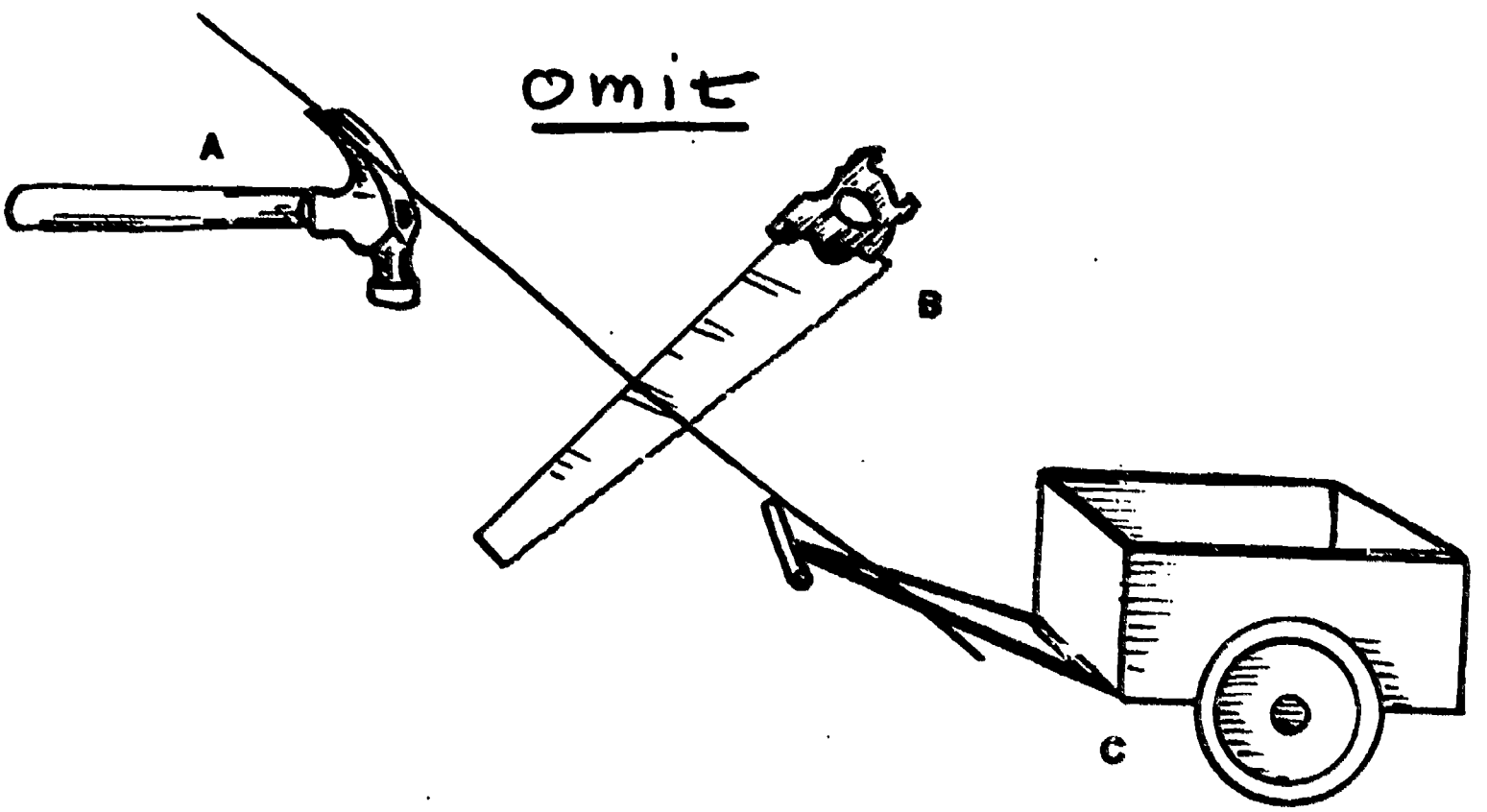




Mag. Simp. Intro. p 5

4. ¿Cuál de las siguientes es un hacha?

omit



EVALUACION DE LAS UNIDADES "MAQUINAS SIMPLES"

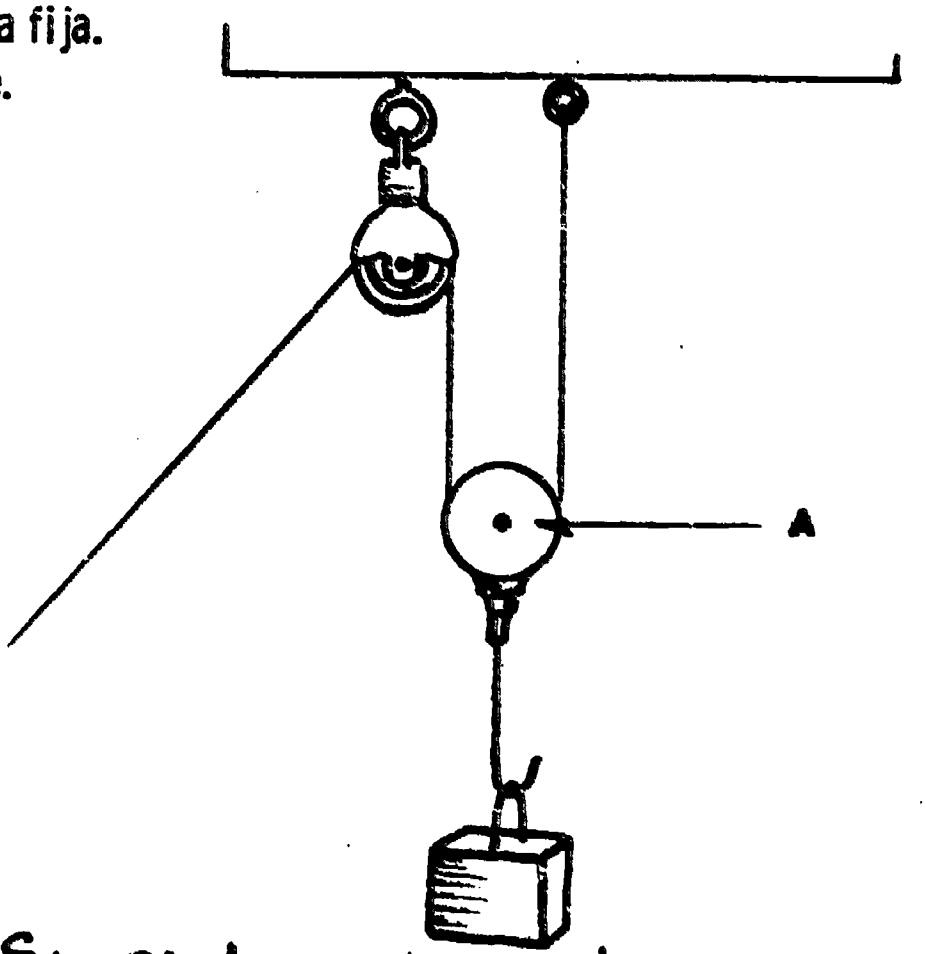
Nombre _____ Fecha _____

Escuela _____ Grado _____

Selección Múltiple (ctda)

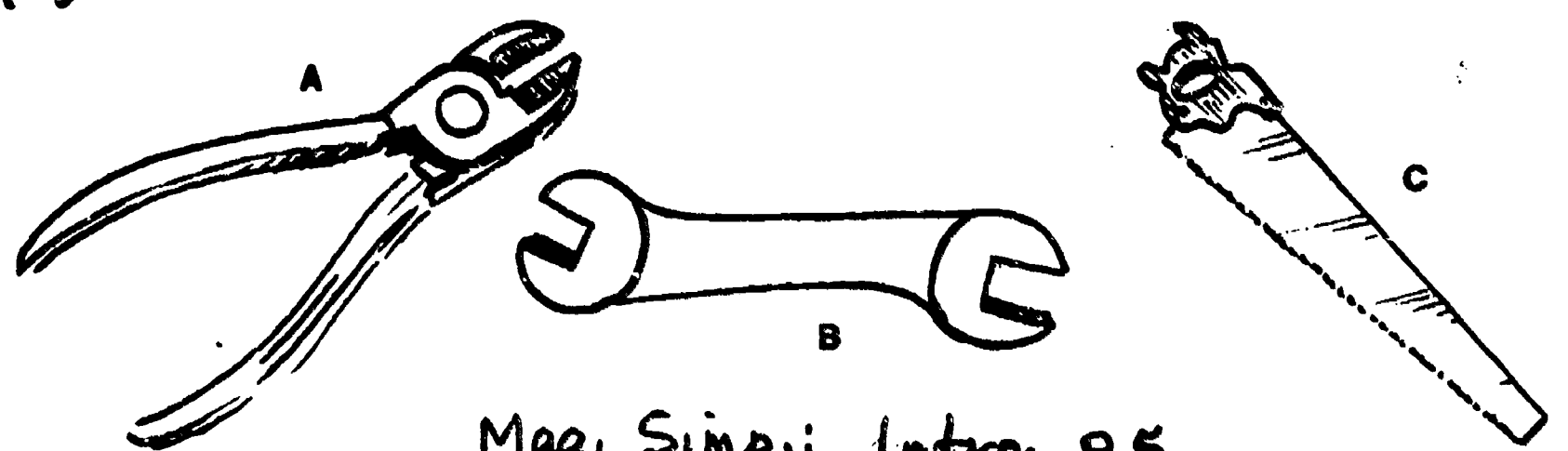
(48) 5. Observa el dibujo. La flecha "A" está señalando

a. el eje de la polea fija.
b. la polea movible.
c. la polea fija.



Maq. Simp: La rueda y polea - p. 10

(90) 6. Para cortar un alambre debe usarse



Maq. Simp.: Intro. p 5.

EVALUACION DE LAS UNIDADES "MAQUINAS SIMPLES"

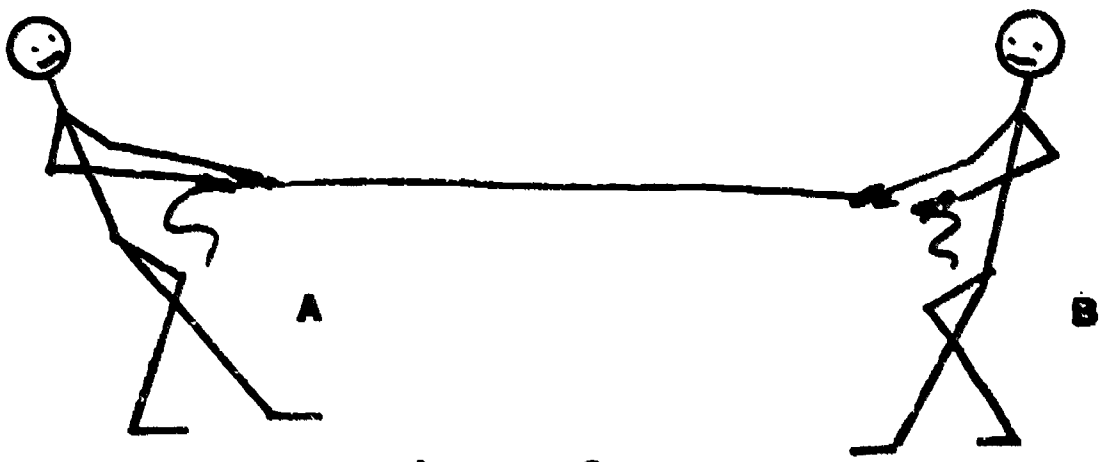
Nombre _____ Fecha _____

Escuela _____ Grado _____

Selección Múltiple (ctda)

(35) 7. Observa el dibujo. Si los muchachos no se mueven de lugar, la fuerza que usa el muchacho "A" es

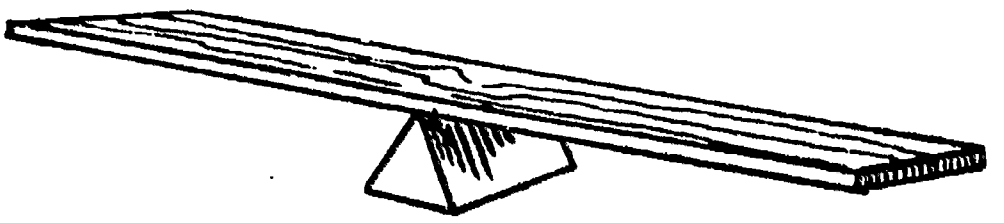
a. igual a la que usa el muchacho "B".
b. mayor que la fuerza hecha por el muchacho "B".
c. menor que la fuerza hecha por el muchacho "B".



Maq. Simp.: Trabajo y Fuerza p 1.

(64) 8. La cuña que sirve de apoyo al tablón se llama:

a. herramienta.
b. fulcro.
c. palanca.



Maq. Simp.: Las palancas p 8

EVALUACION DE UNIDADES DE CIENCIA (EL SISTEMA SOLAR, LA ATMOSFERA, EL TIEMPO)

Nombre _____ Fecha _____

Escuela _____ Grado _____

Selección Múltiple (ctda)

Escoge la mejor contestación y escribe la letra en el espacio a la izquierda:

(68) 9. Lo interesante del Cometa Halley es que

a. sólo puede verse desde la luna.
b. pierde su luz cada diez años.
c. se ve desde la tierra cada 75 años.

Ciencia en Esp.: P. 11

(32) 10. Los astronautas americanos caminaron por primera vez en la luna en

a. 1957.
b. 1969.
c. 1972.

Ciencia en Esp.: 7

(48) 11. Entre Marte y Júpiter hay un sinnúmero de pedazos de uno o varios planetas que probablemente explotaron y que reciben el nombre de

a. satélites.
b. asteróides.
c. galaxias.

Ciencia en Esp.: 7

(42) 12. El planeta más pequeño de nuestro sistema solar es

a. Mercurio.
b. Plutón.
c. Urano.

Ciencia en Esp. P 9

EVALUACION DE UNIDADES DE CIENCIA (EL SISTEMA SOLAR, LA ATMOSFERA, EL TIEMPO)

Nombre_____ Fecha_____

Escuela_____ Grado_____

Selección Múltiple (ctda)

(48) 13. El planeta más cercano a la Tierra es

Ciencia en Esp. - p. 9.

a. Júpiter.
b. Marte.
c. Venus.

(52) 14. El planeta que presenta más similaridad con la Tierra y que sugiere la posibilidad de vida en él es

Ciencia en Esp. p. 9

a. Marte.
b. Saturno.
c. Mercurio.

(72) 15. El más lejano de los planetas es

Ciencia en Esp. p. 11

a. Plutón.
b. Neptuno.
c. Saturno.

(71) 16. El nombre del primer satélite artificial es

Ciencia en Esp. p. 4

a. Gemini.
b. Apolo.
c. Sputnik.

(23) 17. Cuando la Luna tiene su fase iluminada al lado opuesto de la Tierra y no podemos verla, decimos que está en

Ciencia en Esp. p 5

a. cuarto menguante.
b. luna nueva.
c. luna llena.

EVALUACION DE UNIDADES DE CIENCIA (EL SISTEMA SOLAR, LA ATMOSFERA, EL TIEMPO)

Nombre_____ Fecha_____

Escuela_____ Grado_____

(81) 18. En el siglo XVII Galileo inventó

Ciencia En Esp. P. 3

a. la máquina fotográfica.
b. el sistema de radar.
c. el telescopio.

(29) 19. La fuerza gravitacional del sol es 28 veces mayor que la de la Tierra. Un astronauta pesaría allá

Ciencia en Esp. P 3

a. más que en la Tierra.
b. menos que en la Tierra.
c. lo mismo que en la Tierra.

(45) 20. El instrumento que se utiliza para medir la presión del aire se llama

Ciencia en Esp. p. 52

a. termómetro.
b. barómetro.
c. velocímetro.

EVALUACION DE UNIDADES DE CIENCIA (EL SISTEMA SOLAR, LA ATMOSFERA, EL TIEMPO)

Nombre _____ Fecha _____

Escuela _____ Grado _____

Cierto o Falso

Escribe en el espacio en blanco la letra "C" si la oración es cierta o "F" si es falsa:

Percent Correct

(35) 1. Las mareas más altas ocurren cuando el sol y la luna están en línea recta con respecto a la tierra.
Ciencia en Esp. p. 6

(68) 2. La teoría heliocéntrica sostiene que la tierra es el centro del sistema planetario. Ciencia en Esp. p-2

(65) 3. Entre la tierra y la luna hay siempre una distancia de 252,710 millas.
Ciencia en Esp p. 4

(58) 4. El mes sideral tarda 27 1/3 días.
Ciencia en Esp. p5

(65) 5. Las cuatro estaciones son producidas por el movimiento de translación de la tierra. Ciencia en Esp. p. 13

(58) 6. El dióxido de carbono es esencial para el proceso de fotosíntesis.
Ciencia en Esp. p 31

(27) 7. Los tornados son tormentas que se forman únicamente en el agua.
Ciencia en Esp. p. 57

(81) 8. La carretilla de mano es una palanca de segunda clase.
Mag Simp.: Las Palancas p. 9

(39) 9. La polea es una modificación de la rueda.
Mag. Simp.: La Rueda p. 8

(45) 10. El trabajo se mide únicamente en términos de fuerza.
Mag Simp. Trabajo y Fuerza.

EVUALACION DE UNIDADES DE CIENCIA (EL SISTEMA SOLAR, LA ATMOSFERA, EL TIEMPO)

Nombre_____ Fecha_____

Escuela_____ Grado_____

Cierto o Falso

Escribe en el espacio en blanco la letra "C" si la oración es cierta o "F" si es falsa:

Percent corrects

(68) 1. La luna es una estrella porque tiene luz propia. Ciencia en Esp p. 4

(84) 2. En 1957, los rusos lanzaron al espacio el primer satélite artificial. Ciencia en Esp. 2

(52) 3. El movimiento de la luna alrededor de la tierra se conoce como rotación. Ciencia en Esp. p 5

(68) 4. Saturno es el planeta más grande del sistema solar. Ciencia en Esp. P 10

(81) 5. Los cambios metereológicos como lluvia o nieve ocurren en la capa exosférica. Ciencia en Esp. P 25

(55) 6. Los cirros son nubes formadas por cristales de hielo. Ciencia en Esp. P. 52

(88) 7. El lugar donde se apoya una palanca se llama fulcro. Mag. Simp: las Palancas p 3

(48) 8. El clavo es un ejemplo del plano inclinado. Mag. Simp. El plano Inclin. p. 5

(16) 9. La fricción es una fuerza que ayuda a los cuerpos a permanecer en movimiento. Mag. Simp.: La Rueda y Polea p. 5

(42) 10. Si una persona empuja un edificio ha efectuado trabajo. Mag. Simp.: Trabajo y fuerza p 4.

STUDY 4

PRINCIPALS' PERCEPTIONS OF MODEL B

As was stated in the general process evaluation (Study 1), implementation of the MODEL B component has been most difficult. In order to try to gain insights from the problems of this program, which might be useful in the event that similar programs were initiated at other schools, the evaluators interviewd the principals who had had the MODEL B component in their schools.

Procedure

Subjects

All four principals were interviewed. The principal of McKinley School had had a program during 1970-1971. The principal of Bethune School had had an operational program from September 1970 until June 1972. The other two principals had programs which were operational in their schools from September 1970 to the present (3 years); these programs were expected to continue in operation during 1973-1974.

Instrument and Procedure

The senior researcher of the project's evaluation team visited the schools during early June 1973. Interview structure had been worked out to assure that the relevant issues would be discussed and to provide a uniform format for recording the interview results. This is shown in Appendix 4.1. Immediately after each interview, the principal's reactions were recorded on the form.

Analysis

As there were only four principals, it was feasible to rearrange the reactions of the principals by content area and prepare the narrative which follows.

Results

Schools with Operational Programs

Principals of both elementary schools with operational MODEL B programs felt that after a few years of instruction, Spanish-dominant pupils who can handle English sufficiently well should be funneled into mainly English instruction. Both agreed that for pupils who begin first-grade instruction in MODEL B, this funneling should come after the third grade. Both principals felt, however, that Spanish-speaking pupils should continue to receive some instruction in their mother tongue (one hour per

day in Miller, one-half hour per day in Ludlow). They felt that, apart from this minimal contact for Spanish-speaking pupils able to learn in English, Spanish instruction should be given only pupils having difficulty with English.

In other areas, the principals of schools with the two operating MODEL B programs differed in their perceptions. The most telling area of difference involved English-speaking pupils. As was noted earlier in this report, neither of the MODEL B schools had classes of English-dominant pupils coupled with the Spanish-dominant ones as had been the original intent of the MODEL B design. At Ludlow, the principal indicated he felt that the performance of most English-dominant children in his school was too low to have them spend time learning a second language. In contrast, at Miller, the principal stated that the lack of participation by English dominants was a result of his misunderstanding of program goals at the time that he became principal. He stated that, next year, teaming of Anglo and Latino classes would be reinstated as it had been in 1971-1972. He stated that he thought that only selected, talented, English-dominant children should be enrolled in Spanish-as-a-second-language instruction. (A few such pupils have received informal SSL instruction in this manner during 1972-1973, but they have not been considered a part of the LET'S BE AMIGOS program and have not had opportunities to mix with the Spanish-speaking children.)

The two principals also differed in their perception of the services that the project is providing to teachers. The Ludlow principal felt that his contacts with the project director have improved supervision, staffing, and materials to the point where there is no serious problem. The principal at Miller, on the other hand, felt that too much of the monitoring involved observing for the purpose of evaluation but there was insufficient feedback to teachers regarding the problems that had been observed. He also felt that there was a need for more materials. He said that teachers felt that the Miami Materials Package (developed by the Spanish Curriculum Development Center) was inadequate, and that more ESL materials were needed.

In summary, the reactions of the two principals with operating MODEL B programs in their schools reflected a low commitment to bilingualism per se; both preferred to think of their program primarily as a compensatory one for short-term instruction. This was in marked contrast to the intentions of program planners who said that they wished to foster bilingualism as an end in itself and felt that students should have continuous exposure to two languages. On other issues, the principals' views were contrasting, one feeling that the program could be extended to English-dominant pupils, but also feeling that there were problems in management, and the other feeling that the program should serve only Spanish-dominant children, but feeling that most management problems had been handled satisfactorily.

Schools without Operational Programs

The two schools whose programs have been withdrawn were very

different from each other. The principal of the McKinley School stated that there were several problems which prevented the program from operating successfully. He said that programs needed to be tailor-made for each school and that the development of a single program to be used at many sites was a waste of time. He felt that English as a second language was the only desirable program, the one which met both his and his community's needs. He also said that the project directors prevented the program from operating successfully because the program divided the control of teachers (between the principal and the program management) when the teachers should be primarily under the control of the principal.

The Bilingual Program has ceased to operate at the McKinley School but an ESL program continues to serve some Spanish-dominant students.

In contrast, the demise of the program at Bethune School was the result of failure of the Spanish-speaking population to grow as anticipated. The principal stated that in the year the program was begun, a large number of Spanish-dominant pupils entered the school in the elementary grades. In subsequent years, the number of pupils has not continued to grow sufficiently, so that a major bilingual program has not been needed. She stated that she was satisfied with the operation of the program when it was in the school, offered no specific criticisms, and said that she would want a similar program if the need arose. The Bethune principal said also that the teacher assigned to her school during the program was still working with a mixed class, which includes the Spanish-dominant and some English-dominant pupils. He includes ESL and SSL in the instructional program, and makes some use of Spanish in the classroom.

Taken together, the results show a marked contrast between the two schools which have now discontinued MODEL B components. In one school, where there was sharp objection to the emphasis on Spanish, the bilingual elements have disappeared, leaving only ESL. In the other school, some program elements have been maintained, even though there is no formal program.

Evaluators' Comment

With the exception of one school in which the Spanish-speaking population did not grow as anticipated, the results of these discussions with the principals suggest that in the Philadelphia schools the principals' attitudes and expectations for programs are at least as important as the need in the community. It would appear that the idea of studying in a language other than English as a worthwhile goal in itself is not yet accepted among these principals, except where it is needed to meet specific needs of recent inmigrants.

In replicating this program, or in extending the program to new sites, resources should be allocated to assure that principals have a clear understanding of the program and its aims, and agree at least to experiment with a program emphasizing bilingualism as a desired outcome.

APPENDIX 4.1

SAMPLE OF THE STRUCTURED INTERVIEW USED WITH THE PRINCIPALS OF "MODEL B"

School and Principal

Structured Interview of Principals
in Model B

1. Ask Principal to assess the program(s) operating in his school.

Probe for (a) Commitment to bilingualism

for Anglos

for Latinos

2. Ask principal to describe the ideal bilingual program for his school.

(a) Parts of the principal's program which are feasible for short term

(b) Discrepancies between the principal's ideal and the program in operation

3. What were the program strengths and its main problems this year or its last operational year?

Probe about (a) Supervision

(b) Staffing and personnel

(c) Materials and support services

(d) Program content and procedure

(e) Other

4. What kinds of changes need to be made next year?

Study 5

KINDERGARTEN AND PREKINDERGARTEN PUPIL PERFORMANCE IN MODEL A

During the first years of the LET'S BE AMIGOS program kindergarten pupils were examined with the Philadelphia Readiness Test, a locally developed instrument which had been used extensively in the schools to measure number and reading readiness. For use in the LET'S BE AMIGOS program, special Spanish instructions were developed. When pupils were tested on this instrument it was found that, in two of the project's first three years, pupils of both language groups exceeded citywide and earlier Potter-Thomas School performance, and in one year, results were close to or above citywide performance although below the performance of the local Potter-Thomas base line. These base lines were all derived from the 1968 testing, the last citywide administration of the test (See Study 7 Offenberg, 1973).

The prekindergarten pupils have also been tested in the past. The instrument used consisted of a selection of items from the Philadelphia Readiness Test and some new material derived from prekindergarten objectives. Spanish and English instructions were prepared for this test. This test was used to rank pupils in terms of their achievement in prekindergarten, so those with the most skill could enter an enriched all-day kindergarten program.

This study is a partial replication of the work done in the kindergarten assessment as well as tryout of the Boehm Test of Basic Concepts, a published instrument designed to assess a number of basic concepts which were related to stated prekindergarten and kindergarten objectives.

In the kindergarten, pupils were given both the Boehm and the Philadelphia Readiness Test. This permitted (a) comparison of pupils with previous base lines and (b) the computation of a regression line between the Boehm and the Philadelphia Readiness Test, so that program objectives could be restated in terms of the Boehm.

Part 1 of the Boehm was also used to assess children in the prekindergarten in order to determine which should go to the all-day kindergarten program. This change was made after program supervisors agreed that a high score on the Boehm Part 1 was well enough related to program content to be used in this way.

An attempt was also planned to tease out whether the Bilingual Program's all-day kindergarten component was enchancing pupil performance at levels beyond that which could be predicted from earlier prekindergarten testing, but preliminary examination of the number of pupils with regular kindergarten experience who had been in prekindergarten was too low to conduct a meaningful analysis (9 pupils).

In this study only one stated objective was involved, that applying to Philadelphia Readiness Test performance of the kindergarten pupils.

Objective

Kindergarten. Objective 2.7 of the 1969-1970 proposal stated that kindergarten pupils would exceed the citywide mean of 20.1 on the Philadelphia Readiness Test.

Procedures

Program

Teaching, during 90% of the instructional time in prekindergarten and kindergarten, was in the pupils' mother tongue. Both the Boehm test and the Philadelphia Readiness Test were therefore administered in the mother tongue of each pupil taking the test.

The micro-objectives of the instructional program of both the prekindergarten and kindergarten components guided the teachers in teaching size, shape, letter, and number concepts directly related to items appearing on the Philadelphia Readiness Test. In general, teachers were provided with the micro-objective lists, but were not obligated to follow specific instructional procedures. Teachers and supervisors worked out instructional activities which they felt would lead to mastery of concepts which appeared in the lists. The prekindergarten and regular kindergartens met for one-half day.

In addition to these activities, reading in the pupils' mother tongue was introduced in the all-day kindergarten, a program attended primarily by the most able of the previous year's prekindergarten pupils. This class met for the full school day.

The relationship between program content and items on the Boehm was more tenuous. The manual for the test (Boehm, 1970) says that items measure concepts such as space, time, quantity, and some unclassifiable ones as "Touch," "Soft," "Other," and "Without", which are also similar to content of the program. However, there is no direct correspondence between test items and program content.

Evaluation

Instruments. The Philadelphia Readiness Test and its English and Spanish instructions have been described in previous reports of the LET'S BE AMIGOS PROGRAM. (See Offenberg, 1972, Study 7). The test requires that

pupils copy shapes, match words and letters, and perform simple counting and numerical operations. While there is not an exact correspondence between program content and the test, most test items are similar to, or extensions of, micro-objectives specified for the kindergarten level of the Model School programs.

The Boehm Test of Basic Concepts (BTBC) is designed to measure children's mastery of concepts necessary for achievement in the first years of school. The test is read aloud by the teacher, and is appropriate for use with children in three grades--kindergarten, 1 and 2. BTBC results may be used to identify children with deficiencies in these areas and to identify individual concepts in which the pupils could profit from instruction.

The test consists of two parts designed for consecutive administrations, with the items arranged in order of increasing difficulty. This permits use of Part 1 of the instrument to assess skills of the younger, prekindergarten children. The instructions are available in both Spanish and English versions. The standardization process was carried out on the English version, but the publisher claims that tentative work with the Spanish version suggests that the norms for Spanish-speaking children should be about the same as those for English-speaking children.

Administration. The Philadelphia Readiness Test was administered to kindergartens by the regular classroom teachers. Pupils received instructions in their mother tongue. As was the case in the past, teachers administered the tests in small groups, one page of items at a time.

The Boehm test was administered individually by the program supervisor to the prekindergarten pupils (Part 1 only) and to the all-day kindergarten pupils. The Spanish-speaking pupils in the regular kindergarten classes took their tests in small groups while the English-speaking pupils from the same classes took their tests individually. The members of the research staff who did the majority of the testing in these classes selected group rather than individual testing on the grounds of personal preference and comfort with the testing situation. The test publisher suggests that either approach should provide similar results as long as the numbers of pupils tested in groups were small enough to give pupils individualized attention.

On the kindergarten level where both Parts 1 and 2 were given, testing usually took place during two sessions several days apart. Occasionally when a child seemed energetic and enthusiastic after Part 1 and he wanted to continue immediately with Part 2 he was allowed to do so. Testing took place during May and June.

Subjects. All pupils in the kindergarten and prekindergarten programs were tested unless they were absent for extended periods of time. In the kindergarten 157 (69%) of the pupils took both tests, and 32 (14%) of the pupils took one or the other of them. In the prekindergarten 79 (88%) took the Boehm.

Analysis. As the prekindergarten test administration was primarily designed to rank pupils in terms of their readiness for the all-day kindergarten, descriptive statistics were obtained for the part of the Boehm that was applied to this group.

In addition to descriptive statistics, performance of kindergarten pupils on the Boehm Test were compared with the norms for low-socioeconomic-status pupils. Descriptive statistics were also prepared for the Philadelphia Readiness Test. The means obtained were compared with those of the criterion.

In addition, to restate objectives (now in terms of the Philadelphia Readiness Test) in terms of the Boehm, a linear regression equation and correlation coefficient were computed. The equation is for predicting Boehm scores from Philadelphia Readiness Test scores. The correlation coefficient shows the amount of variance that the two tests have in common.

Results

Kindergarten Testing, Philadelphia Readiness Test

Results obtained in the 1973 testing of the kindergarten pupils are in Table 5.1. The mean scores of 22.4 (English-dominant) and 22.6 (Spanish-dominant) show performance very similar (within 0.6 correct items) to that obtained last year, and about one and one-half items more than necessary to meet the objectives. Scores of both groups exceeded those of the base line. As expected, the performance of the all-day kindergarten group exceeded that of the total kindergarten. The 17 English-dominant pupils had an average of 24.1 and the 18 Spanish-dominant children in this group averaged 25.4 (test maximum is 27).

TABLE 5.1. KINDERGARTEN PUPIL PERFORMANCE ON THE PHILADELPHIA READINESS TEST

Item	Pre-program (1968) Base Lines		MODEL A (1973)	
	All Philadelphia	Potter-Thomas	English Dominant	Spanish Dominant
Mean of Group	20.1	20.9	22.4	22.6
Standard Deviation	-	-	5.3	5.9
No. of Children	-	-	95	95

Kindergarten Testing, Boehm Test

Results of the Boehm test administration are shown in Table 5.2. The results obtained were similiar to those on the Philadelphia Readiness Test,

with both groups at about or above the 50th percentile for pupils of their socieconomic status. Results for the all-day kindergarten subgroup showed a high level of performance. Nineteen Spanish-dominant pupils averaged 34.0, (75th percentile) and 18 English-dominant pupils, 32.0 (70th percentile), a result which shows equally superior performance by both groups.

TABLE 5.2. KINDERGARTEN PUPIL PERFORMANCE ON BOEHM TEST

Item	English-Dominant	Spanish-Dominant
Number Tested	87	75
Mean Score	30.6	28.2
Standard Deviation	6.39	6.03
Percentile Rank of Pupil at Mean (Low S.E.S. norms for midyear kindergarten)	65th	50th

Examination of the regression of the Boehm upon the Philadelphia Readiness Test (PRT) showed a moderate correlation of .40 (df=153, p=.01). The regression equation for predicting Boehm scores from PRT scores is

$$\hat{Y}_{(\text{Boehm})} = 13.2 + .70X_{(\text{PRT})}$$

These results were obtained for both ethnic groups combined. On the basis of these results the equivalent of the citywide 1968 mean on the Philadelphia Readiness Test is 27.3, and that of Potter-Thomas in 1968, 27.8.

Prekindergarten Pupils' Performance on the Boehm

The testing of prekindergarten pupils on the first half of the Boehm was designed to answer the question, "Does the difficulty range of Part 1 of the test seem to be appropriate for assigning prekindergarteners to regular

and all-day kindergarten? The results shown on Table 5.3 indicate that the difficulty range is probably satisfactory, with the average child having a mean score of about 60% of the items. There appeared to be little or no meaningful difference between the two language groups on this half of the test when administered to prekindergarteners.

TABLE 5.3. PREKINDERGARTEN PUPIL PERFORMANCE ON BOEHM TEST, PART 1

Item	English-Dominant	Spanish-Dominant
Number Tested	45	34
Mean Score	15.0	15.7
Standard Deviation	4.6	4.1
Percentage of Items Correct	60%	63%

Evaluators' Comments

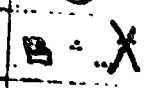
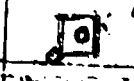
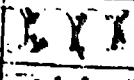
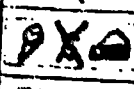
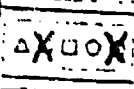
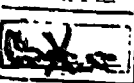
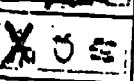
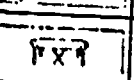
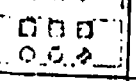
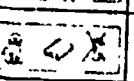
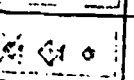
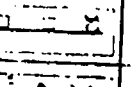
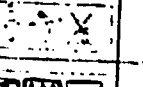
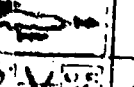
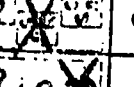
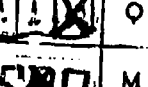
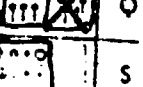
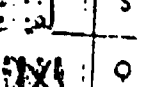
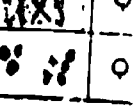
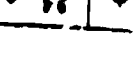
Results of this study were, on the whole, gratifying. It was good to see that pupil performance on the Philadelphia Readiness Test was maintained at a level beyond the minimum stated by the objective despite the interferences which occurred during the school year. It was also pleasing to see that both ethnic subgroups of the regular kindergarten and all-day kindergarten had achieved well in comparison with other children of similar socioeconomic status. The small difference in overall performance of the English- and Spanish-dominant kindergarten groups as wholes on the Boehm test could possibly be explained by several factors. Although the publishers of the test maintain that testing can be done either individually or in small groups without affecting performance, the different grouping in which Anglo and Latino pupils in th regular kindergarten took their tests seems a likely cause of the discrepancy, which was observed only in those classes where the testing situation was different for the two groups. Another factor which might have influenced the scores of regular kindergarten classes was a possible difference in the amount of interaction between items on the Boehm test and the program content in the Spanish versus English regular kindergarten classes. Possible deficiencies in the translation of the test might also have presented more of a problem to the pupils in the regular kindergarten than to the pupils in the more select all-day kindergarten.

To begin to answer these questions, attached to this paper are two appendices. Appendix 5.1 shows the percentage by ethnic group and grade, of pupils who succeeded in completing each item. Appendix 5.2 contains a review of the Spanish translation of the items by a native Spanish speaker which might point to defects in the preparation of this version of the test.

APPENDIX 5.1. PERCENTAGE OF PUPILS SUCCEEDING ON EACH ITEM ON THE BOEHM

Item		Prekindergarten		Reg. Kindergarten		All-Day Kindergarten	
		Eng.-Dom.	Sp.-Dom.	Eng.-Dom.	Sp.-Dom.	Eng.-Dom.	Sp.-Dom.
BOOKLET 1		N =24	N=14	N =69	N=56	N=18	N=19
1. Top	S	100	88	93	89	100	84
2. Through	S	71	62	90	68	94	78
3. Away from	S	60	58	83	71	66	68
4. Next to	S	71	44	86	32	94	84
5. Inside	S	82	85	91	88	94	100
6. Some, not many	Q	58	100	78	88	94	100
7. Middle	S	56	21	78	64	83	79
8. Few	Q	33	53	62	70	50	74
9. Farthest	S	47	79	67	91	89	100
10. Around	S	80	53	97	63	100	84
11. Over	S	67	79	74	93	83	100
12. Widest	Q	33	44	75	52	72	68
13. Most	Q	82	82	94	84	100	100
14. Between	S	44	21	72	32	72	53
15. Whole	Q	56	74	70	91	83	84
16. Nearest	S	96	94	100	96	100	100
17. Second	Q	29	24	45	18	61	47
18. Corner	S	47	38	80	52	61	63
19. Several	Q	87	85	94	86	78	89
20. Behind	S	62	59	72	64	89	95
21. Row	S	42	71	61	82	78	95
22. Different	M	40	53	72	61.	78	58
23. After	T	29	65	59	46	78	89
24. Almost	Q	89	94	84	93	83	100
25. Half	Q	33	38	74	53	72	47

Appendix 5.1. (continued)

Item			Req. Kindergarter Eng.-Dom.	Req. Kindergarter Sp.-Dom.	All-Day Kindergarter Eng.-Dom.	All-Day Kindergarter Sp.-Dom
BOOKLET 2 23. Contur		S				
27. As many		Q	55	80	34	95
28. Side		S	70	48	78	89
29. Beginning		T	65	41	50	74
30. Other		M	48	39	50	26
31. Alike		M	57	39	67	53
32. Not first or last		Q	65	50	78	95
33. Never		T	39	32	55	74
34. Below		S	38	43	55	42
35. Matches		M	35	75	44	94
36. Always		T	52	45	72	79
37. Medium-sized		Q	51	18	44	32
38. Right		S	25	34	44	58
39. Forward		S	44	36	66	11
40. Zero		Q	49	43	33	36
41. Above		S	48	41	55	89
42. Every		Q	58	32	78	68
43. Separated		S	80	45	61	53
44. Left		S	55	41	50	68
45. Pair		Q	33	36	39	47
46. Skip		M	14	14	17	11
47. Equal		Q	26	5	28	21
48. In order		S	17	9	44	89
49. Third		Q	36	17	17	36
50. Least		Q	28	4	11	0
			22	18	17	68

APPENDIX 5.2. A NATIVE SPANISH SPEAKER'S REVIEW OF THE SPANISH VERSION OF THE BOEHM

Miss Leonor Lega, a native Spanish speaker (from Colombia) reviewed the Spanish version of the Boehm. She made the following comments.

"In Latin America there are regionalisms which cause words or phrases which are used in one country to be awkard or misleading in other countries. At least two obvious examples of this problem appear on the test. In a sample item the word platano is used for banana, whereas in most of South America guineo is used. In these areas, platano is a vegetable related to the banana, but served fried.

"The phrase en el mismo medio de is used for right on the...in the test. In Puerto Rican speech, exactamente en would be clearer.

"Both of these instances appear in the instructions and samples of the test and may have caused some pupils confusion in deciding exactly what was to be done.

"Both regionalisms noted seem to reflect the pattern of Mexican speech. Careful review may show that other items have regionalisms characteristic of that country.

"At least one item is poorly translated (No.30). The English version depends on the use of forms of the word other and another: One is an ice-cream, another is a piece of pie. Mark the other (i.e., the third) dessert. The Spanish version uses otro for both other and another, possibly leading the pupil to believe that he is being asked to mark the pie."

STUDY 6

STANDARDIZED TESTING OF SECOND-THROUGH FOURTH-GRADE PUPILS IN THEIR FIRST AND SECOND LANGUAGE

In the third year of the program, comprehensive standardized testing was begun in the MODEL A and MODEL B second and third grades. In this, the fourth year program, the testing conducted previously was expanded to include fourth-grade pupils in MODEL A (there were no fourth-grade pupils in MODEL B) and ARRIBA third- and fourth-grade pupils.

In the third-year evaluation report (Offenberg et al., 1973) the following major results of standardized testing at the primary grade levels were cited:

1. English-dominant second-grade pupils in MODEL A and MODEL B were shown to have grown beyond pupils in a base line groups by one to six months on each of the four Stanford Achievement Primary Battery II subtests administered.

2. Comparison of the third-grade English-dominant pupils with the pre-program base line showed statistically significant (at least $p < .10$) gains in the Vocabulary, Language, and Composite scores of the Iowa Tests of Basic Skills.

3. Second- and third-grade Spanish-dominant children showed strong gains over a pre-program base line in reading Spanish, as measured by the Test de Destrezas Básicas en Lectura, a Spanish reading test used in Puerto Rico.

4. Third-grade Latino pupils, who had received about one year of instruction in English as a second language, were performing in English like English-speaking children who were at the end of the first grade with grade equivalents ranging from 1.6 (word-study skills) to 1.9 (spelling).

In all of these outcomes, English-dominant pupils were compared with English-dominant pupils who were in the Potter-Thomas School during the last year that the program featured regular, nonbilingual instruction. Spanish-speaking pupils' Spanish reading performance was compared with the results of a citywide testing program conducted in 1968, in which all Spanish-speaking pupils were tested in reading in their mother tongue.

During the 1972-1973 school year the testing which led to the above results was replicated. In addition, pupils in the fourth grade were tested. English-dominant pupils were tested using Level B of the Iowa Tests of Basic Skills, and Spanish-dominant pupils were tested with Prueba de Lectura, a reading test from Puerto Rico which was used in the 1968 Spanish citywide testing.

Fourth-grade pupils in MODEL A and ARRIBA were also given the tests normally used in the second or third grade with native speakers of

their second language. Thus, the English-dominant pupils were tested with the Test de Destrezas Basicas, and the Spanish-dominant second graders were tested with the Stanford Primary Battery II. Program planners' pupil-performance expectations in second language were clearly stated for MODEL A and MODEL B Spanish-dominant children: third-grade pupils should be about 1.8, and fourth-grade pupils should be about 2.8 on norms for native English-speaking pupils. There were no clear expectations for the English-dominant group.

In the ARRIBA program, pupils were generally recent migrants. It was therefore felt that testing in Spanish was appropriate for this group, but testing in English might not be. After considerable discussion, it was decided that pupils in the ARRIBA third and fourth grades would be tested in English as well as Spanish, more to obtain an idea of where pupils were than to assess the program.

Procedure

Program

MODEL A instruction was provided by teams of two teachers, one English-dominant and one Spanish-dominant. At the lowest grade levels both teachers in the team knew English and Spanish. In the upper grades some of the English-dominant teachers did not speak Spanish. While in theory teams taught one grade level (or had a mixed class of one or two grade levels, pupils went to instructional groups consisting of pupils who were at approximately the same instructional levels rather than being kept with students in the same grade level. Pupils in the program received instruction primarily in their mother tongue with only about 10% in their second language. In the upper grades (3rd and 4th), instruction was half a day in each language.

In MODEL B, instruction was given in grade-level classes. All pupils involved were Spanish-dominant and teaching teams included Spanish-dominant and English-dominant teachers. Otherwise, it paralleled Model A.

In the ARRIBA third and fourth grades, instruction was provided by two teachers, one English-dominant and one Spanish-dominant. Because the program is designed for pupils newly arrived from Spanish-speaking areas, about 4/5 of the day's instruction was in Spanish (Spanish as a first language, science, social studies, and mathematics). The Spanish instruction paralleled the regular Philadelphia curriculum in these subject areas as closely as possible. In addition all pupils received English-as-a-second-language instruction which began primarily with oral work, but gradually introduced reading in English as well.

The Laidlaw Reading Series was used for Spanish instruction; the Bank Street Readers were used for English instruction in Reading Levels 1 to 3; the Lippincott Readers were used for Level 4 and above in English.

Evaluation

During May and June, standardized tests were administered to the pupils of the program. In MODEL A tests in the pupils' mother tongues were administered by classroom teachers. During the first-language testing, members of the research staff walked from room to room to assure that teachers administered the tests according to published instructions. Second-language testing was administered to pupils in the auditorium, a room which has provision for writing. There, all pupils taking a test were tested at once.

Testing in both languages was conducted in several sessions of one hour or less to assure that students were not overtired by the procedures.

In MODEL B and ARRIBA, tests were administered in the classroom by the teachers. In virtually all the testing sessions a member of the evaluation staff was present to monitor and to assist the teachers.

Instruments. The tests used were as follows:

.Stanford Achievement Test, Primary Battery II - four subtests Word Meaning, Paragraph Meaning, Spelling, and Word-Study Skills. This test was administered to English-dominant second graders and to third- and fourth-grade Spanish-dominant children.

.Iowa Tests of Basic Skills, Level A was administered to third-grade English-dominant pupils, Level B to fourth-grade English-dominant pupils.

.Test de Destrezas Basicas en Lectura (test of Basic Reading Skills). This was administered to Spanish-dominant second- and third-grade pupils, and to fourth-grade English-dominant pupils.

.Prueba de Lectura (Reading Test). This was administ red to fourth-grade Spanish-dominant pupils.

All testing in the mother tongues was done with instruments prescribed for the pupils' actual grade level. Tests used for second language were for levels of one to two years younger than the pupils examined with them.

In administering the tests to pupils in their first language, publishers' directions were followed explicitly. In the second language, instructions were translated into the pupil's mother tongue after first being read in the target language of the test.

Subjects. Due to pressures resulting from the school strike and the need to give both a first-language and a second-language battery to many pupils, makeup tests could not be provided. As a result, numbers of tests administered were frequently lower than the number on roll. This was especially a problem among pupils who took the Iowa Test. Administration took several days, and the data analysis program dropped all pupils

who had missed any test part. In addition, Spanish-language testing was confined to pupils who had enrolled in the Bilingual Program in or before September 1972. The following list indicates the number of pupils on roll in April 1973:

.MODEL A: 248 second grade, 116 third grade, 192 fourth grade.

.MODEL B: 51 second grade, 40 third grade.

.ARRIBA: 15 third grade, 25 fourth grade.

Analysis

Results of testing pupils in their mother tongue were compared with historical base lines. The Spanish-language base lines were from the April-May 1968 testing of Spanish-speaking pupils in the city of Philadelphia. English-language performance was compared with performance of English-dominant pupils at the Potter-Thomas School the last time the tests were given before the Bilingual Program was instituted in the grade: 1970 for the second grade, 1971 for the third grade, 1972 for the fourth grade. During 1971 and 1972 pupils in the base line groups were all participating in a special reading program being carried out in the entire "District 5" area in which the Potter-Thomas School is located and thus may have scores that have been elevated by this districtwide program.

In the base line group, some fourth-grade pupils were tested out of level (i.e., they took Iowa test Level A). As a result, grade-equivalent scores were used in the Iowa test analyses.

Analysis of variance or multivariate analysis of variance (when subtests were analyzed) were used to compare pupils in the program with the base line groups. These analyses were of the grade equivalents (for the Iowa tests) or raw scores (all other tests).

In second language, where no base line was available, descriptive statistics (mean, standard deviation, percentile rank of the mean, etc.) were provided. Descriptive statistics were also provided for the tests of performance in the ARRIBA program because there was no appropriate base line group with which these pupils could be compared.

In addition to these analyses an attempt was made to see if fourth-grade MODEL A pupil performance in the second language was dependent upon length of experience in the program. English-dominant pupils' scores were correlated with the length of time that the pupils had been in the program (one, two, or three or more years).

The situation was more complex for Spanish-dominant pupils because of the importance of the use of the language in the home. As a result, partial correlation was used in which the effect of length of experience in the program was observed when family language (Spanish only or bilingual) and birthplace (Spanish-speaking area or mainland of the United States) were partialed out.

Results

Spanish-Speaking Pupils--Mother Tongue

Second- and third-grade Spanish-dominant pupils were tested with the Test de Destrezas Basicas En Lectura, and compared with the preprogram 1968 base line. Results in Table 6.1 showed that the average child in MODEL A and MODEL B, in both the second and third grades, was between the 60th and 65th percentiles.

TABLE 6.1 RESULTS OF TESTING SPANISH-DOMINANT PUPILS ON THE TEST DE DESTREZAS BASICAS EN LECTURA

Item	Model A	Model B	1968 Base Line
Second Grade:			
Mean Score	79.7**	80.2**	60.8
Percentile Rank*	60.0	62.0	35.0
Standard Deviation	26.7	18.1	16.2
Number of Pupils	81.0	25.0	366.0
Third Grade:			
Mean Score	95.0**	93.1**	70.2
Percentile Rank*	65.0	60.0	27.0
Standard Deviation	19.7	18.6	26.7
Number of Pupils	72.0	27.0	366.0

Analysis of Variance

	MS	F	df	p
Program Treatment	36608.4	118.7	2	.001
Grade	26235.4	85.1	1	.001
Interaction	584.1	1.9	2	NS
Within Cells	308.4		932	

*Norms are for the end of the year, rural zone of Puerto Rico.
**Differs from 1968 base line to a statistically significant ($p < .001$) degree.

This compared with performance at the 35th and 27th percentiles in the pre-program base line. Analysis of variance showed a clear-cut significant difference between groups, a difference which was obviously attributable to the superiority of performance of pupils in both MODEL A and MODEL B when compared with the base line groups. Strictly speaking, further analysis would normally be required to show that the difference between groups was attributable to the differences between the base line and the other groups. However, the mean differences were so clear-cut as to make further analysis appear redundant.

At one ARRIBA school, Waring, a few third-grade students were enrolled in the program. The 10 pupils who completed the Test de Destrezas Basicas had an average score of 96.0 (standard deviation was 15.7). The percentile score of this average was 65.

Prueba De Lectura

Performance of MODEL A, Spanish-dominant pupils was compared with the 1968 base line. The results, in Table 6.2, showed that the program had substantially improved pupil performance.

Table 6.2 PERFORMANCE OF SPANISH-DOMINANT FOURTH-GRADE PUPILS (IN MODEL A) ON THE PRUEBA DE LECTURA

Item	1968 Base Line	Model A
Mean Score	17.7	31.3*
Percentile Rank	5.0	40.0
Standard Deviation	8.7	23.3
Number of Pupils	180.0	74.0

Analysis of Variance

	MS	df	F	P
Program	9686.7	1	45.8	.001
Within Groups	211.0	252		

*Differs from 1968 base line to a statistically significant ($p < .001$) degree.

MODEL A pupils were at the 40th percentile as compared to the 5th- percentile performance of the base line group.

In addition to the MODEL A group, 15 pupils in the ARRIBA program (4th grade) completed the Prueba de Lectura. The mean score for these pupils was 13.6, equivalent to a score at the fourth percentile. The standard deviation was very large, 11.1 (the distribution of scores was skewed: the highest raw score was 45; the lowest was zero).

English-Language Performance, English-Dominant Children in MODEL A

Stanford Achievement Test subtests were used to assess the performance of pupils in the second grade. Results of this testing, including comparison with the Spring 1970 base line, are shown in Table 6.3.

TABLE 6.3. PERFORMANCE OF ENGLISH-DOMINANT SECOND-GRADE "MODEL A" PUPILS ON THE STANFORD ACHIEVEMENT PRIMARY BATTERY II

Item	Word Meaning	Par. Meaning	Spelling	Word-Study Skills
Model A Program (N=63)				
Mean Score	9.7*	11.9*	6.8*	24.7*
Grade Equivalent	1.9	1.7	2.3	1.9
Standard Deviation	5.6	7.5	5.6	7.5
Base Line (1970) (N=42)				
Mean Score	6.0	8.6	2.1	16.5
Grade Equivalent	1.7	1.6	1.5	1.4
Standard Deviation	3.8	6.3	2.1	6.3

	F	df	P
Multivariate F	10.8	4/100	.001
Word Meaning	13.7	1/103	.02
Paragraph Meaning	5.7	1/103	.001
Spelling	27.5	1/103	.001
Word Study Skills	34.7	1/103	.001

*Differs from 1970 base line to a statistically significant ($p < .02$ or better) degree.

The multivariate analysis of variance showed that all of the differences between the program and base line groups were statistically significant. Examination of the grade equivalents of the means showed that program pupils were two months ahead of base-line pupils in Word Meaning, one month ahead in Paragraph Meaning, eight months ahead in Spelling, and five months ahead in Word-Study Skills.

Iowa Tests of Basic Skills

Third- and fourth-grade English-dominant pupils were tested with the Iowa Tests of Basic Skills. (Level A for the third grade, Level B for the fourth grade). Results in Table 6.4 showed that, despite the disruptions in the school year performance of English-dominant pupils was, for the most part, similar to that of the base-line group.

As expected, the multivariate analysis of variance revealed clear-cut differences in performance between third- and fourth-grade pupils on all four measures in the base line and program groups, with third-graders' grade equivalents around 2.5 on all subjects and fourth-graders' performance around 3.3.

Examination of the four subtests to see if the program produced differences showed that only the Arithmetic difference approached statistical significance ($p < .15$). In the third grade, pupils in the program exceeded the base line by about two months of GE score; in the fourth grade, by about one month of GE score.

The apparent effect of the program on pupil performance in Language Usage showed a complex situation: the significant interaction for this variable suggested that pupil performance in the program was ahead of that in the base line in the third grade, but below that in the fourth grade.

In Reading, no clearly significant differences were observed at all.

In Vocabulary, the marginal main effect ($p < .15$) for program and the interaction between program and grade showed that there was a weak tendency for fourth-grade pupils to be slightly behind the base-line group while third graders remained essentially the same as the base-line group.

Given the inconsistency of these effects and the lack of any really significant main effect for treatment, the most parsimonious conclusion is that there was little difference between performance of English-speaking pupils in the base-line and program groups at the times when they were tested.

TABLE 6.4. PERFORMANCE OF ENGLISH-DOMINANT PUPILS ON IOWA TESTS OF BASIC SKILLS

Item		Vocabulary	Reading	Language	Arithmetic
Grade 3					
Program	Mean GE	2.34	2.49	2.56	2.72
(N=42)	SD	.63	.57	.50	.72
Base-line	Mean GE	2.32	2.35	2.33	2.54
(N=86)	SD	.59	.61	.53	.69
Grade 4					
Program	Mean GE	3.00	3.11	3.05	3.47
(N=56	SD	.71	.86	.72	.80
Base-line	Mean GE	3.36	3.23	3.39	3.32
(N=73)	SD	1.46	1.19	1.03	.94

Multivariate Analysis of Variance

	F	df	p
Grade			
Multivariate	23.4	4/250	.001
Vocal	55.8	1/253	.001
Reading	53.2	1/253	.001
Language	81.9	1/253	.001
Arithmetic	62.0	1/253	.001
Program			
Multivariate	2.67	4/250	.03
Vocal	2.23	1/253	NS**
Reading	0.001	1/253	NS
Language	0.53	1/253	NS
Arithmetic	2.49	1/253	NS**
Interaction			
Multivariate	3.01	4/250	.02
Vocabulary	2.32	4/253	NS**
Reading	1.47	1/253	NS
Language	8.49	1/253	.004
Arithmetic	0.02	1/153	NS

*Within grade levels, mean scores on respective subtests in 1973 did not differ from 1968 base lines to a statistically significant degree.
**p <.15.

TABLE 6.5. ENGLISH PERFORMANCE OF SPANISH-SPEAKING STUDENTS ON THE STANFORD ACHIEVEMENT TEST.*

Group	Word Meaning			Paragraph Meaning			Spelling			Word-Study Skills			Composite
	Mean	N	SD	Mean	N	SD	Mean	N	SD	Mean	N	SD	Mean
MODEL A													
3rd Grade	2.20	33	.46	1.96	43	.55	2.04	30	.87	2.15	40	1.06	2.08
4th Grade	2.18	75	.58	2.04	80	.55	2.14	70	.98	2.16	75	.87	2.12
MODEL B													
3rd Grade	2.14	19	.51	1.83	19	.39	2.07	25	.71	1.58	21	1.43	1.90
ARRIBA													
3rd Grade**	1.98	10	.40	1.92	10	.57	1.68	9	.63	1.50	1[illegible]	.14	1.75
4th Grade	1.77	12	.37	1.75	11	.35	1.54	12	.46	1.41	14	.39	1.60
All Components													
3rd Grade	2.14	62	-	1.92	72	-	2.[illegible]0	64	-	1.87	73	-	1.97
4th Grade	2.12	87	-	2.00	90	-	2.05	82	-	2.04	89	-	2.05

*Grade-equivalent scores.
**ARRIBA third grade is at Waring School; Ludlow has no third-grade program.

Second-Language Performance

Table 6.5 shows the performance of the third- and fourth-grade Spanish dominant pupils in the various components of the program in reading English. Program planners stated that beginning with the second grade, Spanish-dominant pupils in the Model School programs should grow one year for each year of English reading instruction. Thus, third-grade pupils should be reading at about 1.8, fourth-grade pupils should be reading at about 2.8. As can be seen in the table, performance of MODEL A and MODEL B third-grade pupils was at or above the level specified (with the exception of MODEL B performance on the Word-Study Skills subtest). In contrast, the MODEL A fourth-grade group did not reach the desired level of performance on any test.

Pupils in the ARRIBA program were also tested to obtain a rough estimate of their performance in English. This testing was conducted late in June (after the date when school would have ended in a normal year) and many were absent. In addition, the ARRIBA program included many students who had arrived after the beginning of the school year and thus were not eligible to take the test. As a result, the number of pupils taking each test was low.

Second-language test results in ARRIBA were somewhat below those of MODEL A and MODEL B. The ARRIBA composite mean score was 1.75 in the third grade, and even lower 1.60 in the fourth. The low results were anticipated because of the population served by the ARRIBA program at this level--pupils newly arrived from Puerto Rico or other Spanish-language areas.

In order to assess whether English second-language performance in MODEL A was dependent upon the length of time pupils were in the program, Stanford scores were correlated with the number of years pupils had been in the program. Partial correlation was used in which the effect of the degree to which Spanish was used in the home and the effect of the pupil's having been born on the mainland was partially out. The correlation obtained between the number of years and Word Meaning score was .03, for Paragraph Meaning .06, for Spelling .04, and for Word-Study Skills .03. None of the correlations was statistically significant.

Fourth-grade MODEL A English-dominant pupils were tested in their second language, Spanish. The test used was the Test de Destrezas Basicas en Lectura. Performance of this group is shown in Table 6.6. Each subtest is listed separately because the degree to which knowledge of Spanish is necessary varies markedly for the subtests. Letter and Word Recognition do not require real knowledge of Spanish because they involve matching of words and letters and therefore are probably more dependent on visual discrimination than one's knowledge of the language. Word and Paragraph Meaning, on the other hand, require that the children know the language. The average fourth-grade English-dominant pupil was able to read Spanish on these two subtests as well as a native Spanish-speaking third-grade child.

TABLE 6.6. PERFORMANCE OF FOURTH-GRADE MODEL A ENGLISH-DOMINANT* PUPILS IN SPANISH.

Item	Word Recognition	Word Meaning	Comprehension	Composite
Mean Score	62.0	15.6	12.4	91.7
Standard Deviation	4.4	5.6	4.6	15.8
Number of Pupils	59	58	57	56
Percentile Rank**	93	60	60	75

*Nine children in this group were of Spanish background, but were treated as English-dominant in the program. These children knew more Spanish than other English-dominants and received some Spanish instruction with Spanish-dominant pupils.

**Percentile rank based on rural norms for third-grade native Spanish speakers.

Despite the high level of this performance, the English-dominant child's ability to read Spanish was unrelated to the length of time he had been in the program (r=.03).

Evaluators' Comment

Results of this study showed that the program has been very successful in teaching reading in the Spanish language: Spanish-dominant students were at least at the 40th and often at the 60th percentile in their mother tongue. These results confirmed the findings of the 1971-1972 evaluation, that Spanish-speaking pupils' mother-tongue performances have been enhanced greatly over that which was evident in the base-line group. The results also contained a surprise, that the Spanish reading performance of English-dominant students was very good, as good as Spanish-dominant children one year younger.

In contrast, results obtained for English performance were not clear-cut. They showed significant gains for English-dominant second-grade children in their mother tongue, but no clear differences for third- and fourth-grade English speakers in the Program and the Base-line groups. It is not clear how to interpret the lack of difference, but two aspects unrelated to the program might pertain: (a) the lack of gain over the base line might reflect the impact of the strikes; (b) the lack of gain might reflect the fact that the base line already contains some improvement in the performance of English-dominant pupils served by the program due to the "District 5" reading program mentioned earlier.

There was little difference in English performance of Spanish-dominant children from the third to the fourth grade. While the anticipated amount of growth from third to fourth grade was not obtained, comparison of the 1973 fourth-grade results with the 1972 third-grade results is instructive. During the one-year period, the average score increased by four months on the Word Meaning subtest, by two months on the Spelling and Paragraph Meaning subtests, and by five months on the Word-Study Skills subtest. This suggests that the rate of growth during the fourth grade has been heterogeneous and more attention needs to be paid to program development in some skill areas than in others.

A third area which should receive attention is the fourth-grade ARRIBA group. There is no apparent reason for these pupils not to be reading their mother tongue better than the pupils in the base line. This suggests that attention needs to be given to reading instruction at this level of the program.

STUDY 7

STANDARDIZED TESTING OF READING IN THE "ARRIBA" JUNIOR HIGH SCHOOL CLASSES

In 1971-1972 standardized testing of reading in the Spanish language was undertaken for the first time in the ARRIBA junior high schools. Pupils in seventh and eighth grades of the program were compared with a pre-program base line of scores obtained in a 1968 citywide testing of all Philadelphia pupils of Spanish origin (Desing, 1968). Both groups of students were examined using the Spanish version of the Inter-American Test of Reading, L4-CEs. This is the test designated for the junior high school by the test maker.

Results of the 1971-1972 study showed that the mean score in the seventh grade had risen from the 9th to the 25th percentile, and in the eighth grade from the 8th to the 25th percentile. However, the 1972 performance of ninth-grade pupils was at the 9th percentile, with a raw score slightly below that of the eighth graders (there was no ninth-grade pre-program base line).

In the current evaluation (1972-1973) the testing of ARRIBA junior high school students' reading in Spanish has been partially replicated. Due to strikes during the school year, testing was delayed from the end-of-March-beginning-of-April period used in the past to the last week of May, a delay of about two months. This served as partial compensation for the instructional time lost during the strikes, but it did not compensate for the time needed to review material when school resumed after a strike which was nearly as long as a summer vacation.

As a result of these problems, the findings of this study must be accepted as tentative.

In addition to the testing-date change, the 1972-1973 study attempted to ferret out the impact of years of English instruction on performance in Spanish. This was carried out by using information appearing on the pupil information forms to obtain the number of years in which program participants were instructed in English between leaving their mother country and entering the Bilingual Program.

Procedures

Program

The ARRIBA component was designed to serve Spanish-speaking pupils who had come to the Philadelphia schools directly from a Spanish area and other pupils whose performance in regular English classes suggested that they could be helped by instruction conducted in Spanish. At the junior high school level, Spanish-as-a-first-language courses were required for all participants in the program, as well as other courses in Spanish and ESL. Students who had been in the ARRIBA program before junior high school received specific instruction in reading in the Spanish language. In general, elementary school Spanish reading was taught using the Laidlaw Reading Series. By the junior high school, these texts were often replaced with teacher-selected materials designed to meet the more mature interests of the pupils.

Instrument

The Inter-American Test of Reading (Level 4, Form CEs) was used to assess Spanish reading skills of junior high school students. Vocabulary Level of Comprehension, Speed of Comprehension, and Composite scores were provided. Norms based on an islandwide sampling of pupils in Puerto Rican schools were used.

Subjects

The test was administered to all students in the seventh, eighth and ninth grades who were present in each junior high school on the scheduled testing date in the middle of June. Makeups were given when possible, but many students had ceased attending school because of the summer heat (testing had intentionally been delayed as long as possible because the teachers' strike had so shortened the academic year). One hundred seventeen pupils were tested. This was 67% of the pupils on roll in April.

Method

Testing took place in one session. At Penn Treaty, the students were gathered in the cafeteria and directions were given by a Spanish-speaking teacher. At Stoddart-Fleisher, the students remained in their homerooms and directions were given by a Spanish-speaking

member of the research staff. In both schools monitoring was done by both teachers and members of the research staff.

Base line data from a study (Desing, 1968) conducted before the institution of the Bilingual Program were available for comparison with scores obtained by seventh and eighth graders in 1972-1973. In 1968, Spanish-speaking students in the Philadelphia schools through grade eight took tests which examined Spanish reading ability. The Inter-American Test, Level 4, CEs was taken by seventh and eighth graders and the results obtained in that testing were used to determine the progress which has been made by students in the Bilingual Program.

Analysis

Analysis of variance of raw scores was used to determine whether pupil performance in the 7th and 8th grades was superior to that of the pre-program base line group.

The mean and standard deviation were computed for the ninth-grade pupils because there was no pre-program base line for this grade level.

A Pierson product-moment correlation coefficient was also computed between the raw scores obtained on the test and the number of years in which pupils were involved in English instruction. This analysis was confined to the ninth-grade group.

Results

Seventh and Eighth Grades

Overall performance of seventh- and eighth-grade pupils on the Inter-American test is shown in Table 7.1. Pupils in the program earned statistically significantly better scores than did students before the program. However, pupils did not perform as well as had the similar pupils in the equivalent grade levels in 1971-1972.

Ninth Grade

The 37 ninth-grade pupils had an average raw score of 39.3, and a standard deviation of 13.3, which put the average program participant at about the 30th percentile. This was a clear gain over the

previous year's results at the 9th percentile.

TABLE 7.1. PERFORMANCE ON INTER-AMERICAN READING TEST IN SPANISH (L4-CEs)*

Item	Base line Group 1968	ARRIBA Group 1973
Grade 7		
Number of Pupils	97	41
Mean Score	19.1	25.3**
Standard Deviation	12.3	11.2
Percentile of Mean	9	21
Grade 8		
Number of Pupils	81	39
Mean Score	23.2	29.2**
Standard Deviation	14.9	12.6
Percentile of Mean	8	16

Analysis of Variance

Source	Analysis Ms	df	F	p
Treatment	2159.4	1	72.6	.001
Grade	1047.3	1	6.1	.02
Interaction of Treat & Grade	0.692	1	.004	NS
within groups	171.5	254		

*Spring Semester norms were used (p.53 of the Test Manual). As base line data were grouped in 10-point intervals, the 1973 data were similarly grouped, and analysis was based on the grouped data. This resulted in slight differences from the means and standard deviations appearing in last year's reports, for which ungrouped data were extracted from the original answer sheets.

**Differs from base line to a statistically significant (p=.001) degree.

Relation of Performance to Amount of English Instruction

An attempt was made to determine whether English instruction in regular classes breaking the continuity of Spanish instruction was detrimental to the Spanish-reading preformance of program participants. To this end, a Pierson product-moment correlation coefficient was obtained between the students' test scores and the number of years they had received instruction conducted in English. For all students for whom data were available, a correlation of -.17 (df=29) was obtained, indicating a very weak and statistically insignificant tendency for pupils with a year or more of instruction English to do more poorly on the Inter-American test than pupils without the English instruction. However, it was found that about two-thirds of the pupils had had continuous Spanish instruction, a fact which probably would keep the correlation low, even if there was a relationship.

Evaluators' Comment

Overall, results continue to show that the scores of seventh- and eighth-grade pupils have improved from a base line score attributable solely to guessing, to a score which, though below grade-level norms, represents an ability to read the test and do some work on it. There was a small decline in scores from 1972-1973 for seventh- and eighth-grade pupils, as expected because of the strike.

In contrast to these small declines, the ninth-grade pupils performed better in 1973 than they had the previous year. As 59% of the pupils in the ninth-grade group had participated in the program in the previous year, this gain may reflect the cumulative effect of program participation.

STUDY 8

EFFECT OF PARTICIPATION IN THE "ARRIBA" PROGRAM IN TENTH GRADE ON GRADUATION PREVALENCE IN TWELFTH GRADE

One major objective of the ARRIBA component was to reduce dropout incidence of pupils in its high school level. The last two evaluation reports (Offenberg, 1972, 1973) showed that the program was effective in preventing pupils from dropping out during the school year.

In both years, the percentage of program participants dropping out between October and March was computed. The results of these computations were compared with parallel data for Spanish-speaking pupils attending Edison, Kensington, Franklin, and Penn high schools in 1969, the year prior to the beginning of the Bilingual Program. Results were very similar in both years. For example, in 1971-1972 the dropout incidence was reduced from 25.1% to 8.5% in the tenth grade and from 22.7% to 9.5% in the eleventh grade. The dropout rate showed little change in the twelfth grade, from a low 6% to 8%. The pattern in these results suggested that the program was effective in eliminating most tenth- and eleventh grade dropouts during the time when school was in session for pupils who were enrolled in ARRIBA program offerings.

This analysis might be criticized for two reasons. First, dropout-incidence reductions might be maintained only during the period when pupils were actually receiving instruction in the ARRIBA component. If this were so, the dropout incidence might rise sharply for pupils during subsequent years if they had no contact with the program.

Secondly, dropout reductions might occur during those parts of the year examined in the previous studies (October to March) only to increase in the parts of the year not examined. For example, many participants discouraged from leaving school by the program might withdraw during the late spring and summer hiatus.

The real function of computing dropout rates was to show whether the ARRIBA program has increased the probability that a student will graduate from school. If it were known that a pupil has graduated, it would also be known that he did not leave over vacation intervals or during the years before graduation which followed his participation in the program.

Since 1970, a student-by-student record has been kept for program participants. This record permitted the project to answer the question of the effect of participation in the ARRIBA program three years ago on the actual proportion of pupils who graduated from school in the current year. This study addressed itself to answering that question.

Procedure

Program

The ARRIBA component high school program remained basically unchanged over the three years during which the students examined in this study were in school. Course work was offered in ESL and Spanish work in Spanish as a first language, science (including biology and chemistry), social studies, and mathematics (including algebra, geometry). Over the three years when the pupils were in school, course work in commercial subjects in Spanish was added to the Kensington program. Students who were Spanish-dominant or very fluent in Spanish and were enrolled in the four schools in which the program operated had the option of selecting (with their counselors' advice) one or more courses offered by the ARRIBA component each year.

Evaluation

The evaluation was designed to answer the question, "Are Latino students who participate in the ARRIBA program in the 10th grade more likely to graduate three years later than are those who do not participate in the tenth grade?"

Subjects

The subjects of this study consisted of all students listed in the citywide pupil directory as Spanish-speaking in the tenth grade in December 1970. According to the School District's Division of Administrative and Survey Research, if the dominant language in the pupil's home was Spanish, the pupil's name should appear in the file of Spanish-speakers. One hundred sixty-four students appeared in the projects pupil-information file as tenth-grade ARRIBA pupils in 1970. One hundred two (62.2%) of these pupils appeared in the file obtained from the Division of Administrative and Survey Research. (It is not clear whether someone decided that the remaining 37.8% of the program participants were judged by the schools to be from English-dominant homes or were missed through poor information gathering.)

The 102 pupils who appeared in both the project file and the Administrative and Survey Research file constituted the group of subjects in the "program group." All students not listed in the LET'S BE AMIGOS

pupil-information file were treated as controls.

Three possible biases were operating in this comparison group which would tend to make the control group have a higher percentage of pupils graduating than the experimental group if the program had no effect. First, in the high schools in which the program was operating, students who were guided by counselors into ARRIBA classes were generally felt to be those who were having difficulty in performing in regular, English-speaking classes. The result of this procedure was a bias in favor of control-group students in schools where ARRIBA was offered. Secondly, students were classified on the basis of their participation in ARRIBA as tenth graders. Among those pupils in the program were some who could have chosen not to participate in subsequent grades. This crossing over of pupils was likely to result in underestimation of the true impact of the program. Thirdly, as the ARRIBA component serves high schools with very low socioeconomic status, (Sixty-one percent of the families of pupils at Penn, 50% at Franklin, 47% at Kensington, and 39% at Edison receive Aid to Families of Dependent Children. These schools rank 1, 3, 4, and 7 of the 27 high schools in the city in terms of the proportion of families receiving AFDC.) Spanish-dominant children outside the ARRIBA schools are of higher socioeconomic status. Control-group pupils would then be from areas where the anticipated graduation rate would be higher.

While there was no clean way of eliminating any of the three problems, making two different sets of comparisons would change the impact of each of the problems. In the first set of comparisons, program students were compared with Spanish-dominant students from all over the city. This reduced the impact of the first and second biases above at the expense of adding the third. Comparison of program students with Spanish-dominant students at only their own school eliminated the third bias but increased the impact of the first and second. To provide as fair a representation of program outcome as possible, both control-group comparisons were made. In compiling the data for this study, the 102 students in the ARRIBA program were compared with the 326 Spanish-speaking students in the pupil directory but not in the ARRIBA program, on a citywide basis, and also with the 196 Spanish pupils not participating in the program who were enrolled in tenth grade in schools served by the program.

Method

The current pupil directory was used to locate any pupils who had changed schools due to moving. Graduation lists were then checked to determine whether each student had graduated in the Spring of 1973.

Analysis

Percentage of pupils graduating was computed for each group, and chi-squares were computed to determine the statistical significance of the differences observed.

Results

Table 8.1 compares tenth-grade program pupils with Spanish-dominant pupils citywide. This is the analysis in which the control-group was biased by high socioeconomic status. As can be seen in the table, despite the possibility of this bias, there were strong, statistically significant differences in favor of the program, with participants 1.8 times as likely to graduate as the base line group.

TABLE 8.1 GRADUATION RATES OF SPANISH-DOMINANT PUPILS GRADE 10 IN DECEMBER 1970

Group	Graduated		Not Graduated		Total	
	N	%	N	%	N	%
In ARRIBA	38	(37.3)	64	(62.7)	102	(100)
Not in ARRIBA (Citywide)	68	(20.9)	258	(79.1)	326	(100)
Chi-square = 10.3, df = 1, $p < .005$						

Comparison of students in the program with others in their own schools is shown in Table 8.2. This comparison was probably biased by selection of pupils for the program: pupils whose counselors thought would have difficulty in English classes were more likely to be enrolled in this program than pupils who were doing well in school. As can be seen in this table, the statistically significant results were even more dramatic in this set of comparisons: program students were more than three times as likely to graduate as were other Spanish-dominant students in the same school.

From these figures it was easy to calculate the graduation rate of pupils in schools other than those served by the program--schools with a much higher socioeconomic status and schools offering special vocational and technical programs. The result was 57%.

TABLE 8.2. GRADUATION RATES OF SPANISH-DOMINANT PUPILS IN GRADE 10 IN DECEMBER 1970

Group	Graduated		Not Graduated		Total	
	N	%	N	%	N	%
In ARRIBA	38	(37.3)	64	(62.7)	102	(100)
Not in ARRIBA (Same School)	21	(10.7)	175	(89.3)	196	(100)

Chi-square = 57.4, df = 1, p<.001

Evaluators' Comment

The results of the study confirm previous evidence that pupils enrolled in the Bilingual Program are completing school at a higher rate than are similar students in regular classes.

The large discrepancy between the rates for ARRIBA and non-ARRIBA pupils at the four high schools served is especially deserving of attention. Given the selection procedure used (counselors' recommendations) to enroll students in ARRIBA and the bias resulting in

favor of nonprogram students, discrepancy was initially a surprise. However, consideration of a finding in last year's evaluation may explain this outcome. In a follow-up of pupils enrolled in regular classes (Offenberg, 1973) it appeared that there might be a discrepancy between the amount of English needed for a counselor to recommend that a pupil be enrolled in a regular English-language classroom and the amount of English needed to succeed in that classroom. The large discrepancy between dropout rates of ARRIBA students and of other Spanish-dominant students in the same schools might be attributed to a similar phenomenon.

BIBLIOGRAPHY

Bank Street College of Education. Bank Street Reading Series. New York: Macmillan, 1968.

Boehm, A. E. Boehm Test of Basic Concepts. New York: The Psychological Corporation, 1969.

Departamento de Instruccion Publica. Test de Destrezas Basicas en Lectura. Hato Rey, Puerto Rico: Departamento de Instruccion Publica, 1963.

Departamento de Instruccion Publica. Prueba de Lectura. Hato Rey, Puerto Rico: Departamento de Instruccion Publica, 1966.

Kelly, T., Madden, R., et al. Stanford Achievement Test, Primary Battery II. New York: Harcourt Brace Jovanovich, 1965.

Lindquist, E. F., Hieronymus, A. N., et al. Iowa Tests of Basic Skills. New York: Houghton Mifflin Co., 1955.

Manuel, H. T. Inter-American Test of Reading. Austin, Texas: Guidance Testing Associates, 1962.

McCracken, G., and Walcutt, Basic Reading. Philadelphia and New York: Lippincott, 1965.

Offenberg, R., et al. Title VII Bilingual Project Let's Be Amigos: Evaluation of the First Year, 1969-1970. Philadelphia: School District of Philadelphia, 1970. ERIC Document ED 046-295.

Offenberg, R., et al. Title VII Bilingual Project Let's Be Amigos: Evaluation of the Second Year, 1970-1971. Philadelphia: School District of Philadelphia, 1972. ERIC ED 069-158.

Offenberg, R., et al. Title VII Bilingual Project Let's Be Amigos: Evaluation of the Third Year, 1971-1972. Philadelphia: School District of Philadelphia, 1973. ERIC in press.

Pastore, E., Guzman, R., and Gomez, C. Por el Mundo del Cuento y la Adventura. River Forrest, Illinois: Laidlaw Brothers, 1969.

School District of Philadelphia. Levels in Mathematics for the Primary Years. Philadelphia: School District of Philadelphia, 1968.

School District of Philadelphia. Proposal for the Let's Be Amigos Project. Philadelphia: School District of Philadelphia, 1969.

School District of Philadelphia. Continuation Proposal for the Let's Be Amigos Project, 1970-1971. Philadelphia: School District of Philadelphia, 1970.

School District of Philadelphia. Continuation Proposal for the Let's Be Amigos Project, 1971-1972. Philadelphia: School District of Philadelphia, 1971.

School District of Philadelphia. Continuation Proposal for the Let's Be Amigos Project, 1972-1973. Philadelphia: School District of Philadelphia, 1972.

School District of Philadelphia. Continuation Proposal for the Let's Be Amigos Project, 1973-19 Philadelphia: School District of Phila. delphia, 1973.

Spanish Curriculum Development Center. Miami Curriculum Materials. Miami Beach, Florida: Dade County Public Schools, 1971.