DOCUMENT RESUME

ED 038 238 24 RC 004 272

TITLE	Rock Program Evaluation, 1968 - 1969.
INSTITUTION	Southwestern Cooperative Educational Lab., Albuquerque, N. Mex.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.; Region One Educational Service Center, Edinburg, Tex.
BUREAU NO	BR-6-2827
PUB DATE	10 Apr 70
NOTE	55p.
EDRS PRICE	EDRS Price MF-$0.25 HC-$2.85
DESCRIPTORS	Diagnostic Tests, *English (Second Language), Mexican Americans, Migrant Education, Non English Speaking, *Oral English, *Preschool Tests, *Program Evaluation, Reading Instruction, Reading Programs, Reading Tests, *Spanish Speaking, Testing

ABSTRACT

The Region One Curriculum Kit (ROCK) was utilized as part of a preschool instructional program during the 1968-69 school year. This report presents an assessment of the program. Four groups of Spanish-speaking students were tested by the Michael Test of Oral English Production. Three of the groups utilized the ROCK materials which are designed to increase the oral English development of five-year-old native speakers of Spanish. Group four was a control group and utilized standard basal reading materials. Analysis indicated significant gains in oral language expression for those students in the experimental groups when compared to the control group. The conclusions and recommendations of this report call for an expansion of the program with specific emphasis on teacher training programs. (DK)

ROCK PROGRAM EVALUATION

1968 -- 1969

A report prepared by

SOUTHWESTERN COOPERATIVE EDUCATIONAL LABORATORY, INC.
Albuquerque, New Mexico

for

REGION ONE
EDUCATION SERVICE CENTER
Edinburg, Texas

April 10, 1970

TABLE OF CONTENTS

LIST OF TABLES

BACKGROUND

ESC and Its Role

Region One Education Service Center is one of twenty such Texas institutions keyed to one objective: to make quality ideas, services, information, and teaching materials available to Texas schools whenever and wherever they are needed and desired.

Skilled professionals are prepared to assist educators with cooperative efforts which may range from long-term planning to short-term problems of supply.

Implemented in 1967 by the Texas State Legislature with funds from local, state, and national sources, the Region One Center has developed materials, technology, and consultative help in these areas:

CURRICULUM DEVELOPMENT--Constant updating of what is taught.

DATA PROCESSING--Keeping records and making reports via computer.

DRIVER EDUCATION--Combining techniques of driving with safety habits.

IN-SERVICE TRAINING--New ideas reaching individuals through group effort.

INSTRUCTIONAL TECHNIQUES--Developing new methods of presenting material or combining old ones into more effective teaching packages.

MEDIA SERVICES--Teaching with latest audiovisual materials.

MIGRANT EDUCATION--Tailoring staff development efforts to the needs of teachers whose pupils spend only part of the year in school.

PLANNING--Looking forward and building toward improved educational opportunities within the framework of local, state, and national goals.

PUPIL APPRAISAL--Identifying pupil problems and potential, then developing appropriate follow-up.

SPECIAL EDUCATION--Assisting with specialized instructional approaches for pupils with physical or mental handicaps.

Since more than half of the children entering school in the Region One area speak little or no English, a prime need is to make it possible for them to succeed in learning.

To this end, the Curriculum Division has developed, and is continuing to develop, English as a Second Language and bilingual instructional materials that will help young children learn reading, writing, and social studies in Spanish, as well as in English.

The new curriculum components are being tested and refined so they can be adapted at little cost to schools desiring to implement a bilingual program.

The youngster whose family follows the crops needs a school program especially designed for him if he is to succeed in school.

It is with him in mind that the Region One Migrant Education Division offers its services to educators. Professional consultants

are prepared to conduct summer institutes, as well as year-round workshops in science, language arts, social studies, and reading.

Emphasis is placed on the preparation of the preschool migrant. Counseling is available for every teacher in addition to a comprehensive training program.

The complete kindergarten program, developed by preschool specialists, features English as a second language for five-year olds with teaching aids adapted from Curriculum Division materials.

A viable organization seeks to assess the effectiveness of its programs. This report outlines an extended assessment of the program called "ROCK"--Region One, Curriculum Kit, utilized as a part of the preschool instructional program during the 1968-69 school year.

ROCK Program

The ROCK for preschool is a program for the oral English development of five-year old native speakers of Spanish. It has as its core 128 language lessons written originally at the University of California at Los Angeles and termed the H-200 series. Region One's expansion of the lessons makes possible (1) language instruction in groups of seven children; (2) additional exposure to language patterns under practice by hearing patterns spoken by a variety of voices on audio flashcards; (3) further language development through the use of songs and games written to reinforce the structures being learned and of sound filmstrips of stories, told first in Spanish then in English followed by pattern drills.

Full implementation of the ROCK program requires that each classroom be provided with the kit itself, an audio flashcard machine, and a sound filmstrip projector equipped with headphones. A class is defined as being composed of 20 children (monolingual, multilingual, or a combination of both), a teacher, and a full-time aide.

SWCEL Educational Tie

The Southwestern Cooperative Educational Laboratory (SWCEL) in Albuquerque is one of 15 regional laboratories located in the United States. As a developmental agency which works primarily with non-English speaking populations, it has been enthusiastic about the work under way at the Region One Service Center in Edinburg, particularly that work related to the ROCK materials.

While curriculum materials have been designed for use in classrooms little information about entering abilities of youngsters and performance outcomes has been known. The Michael Test (described later) is one of the few viable alternatives presently known for ascertaining either of the aforementioned imperatives.

A need for a quick-scoring, effective diagnostic instrument--a natural development activity on the part of the Laboratory--and a need for criterion measures led to cooperation between the Region One Education Service Center and the Southwestern Cooperative Educational Laboratory.

Funds for gathering of data, implementation of program materials, and data analyses were granted to Region One by Lee Frasier, Director of the Migrant Education Division, Texas Education Agency.

The educational relationship established called upon the skills and competencies of various organizations in an attempt to provide substantive background data for immediate (when possible) and long-range solutions to long-standing educational problems.

DECISIONS ABOUT THE ASSESSMENT APPROACH

Design

With pragmatic education being a realistic situation, as opposed to the idealistic world of model designers, random assignment of teachers and children to the groups was impossible. Groups were assigned on a geographic basis with various Regional Service Centers serving as the boundaries.

Since randomness was not a fact, the pretest was a necessity with the concept of covariance anticipated to adjust for any unequalities. Introduction of the pre- post-test necessitated a control for test-retest learning. To account for as many threats to validity as possible, an adaptation of the Solomon Four Group Design (reference Gage, Campbell/Stanley, chapter 5) was chosen.

Four various treatment groups were identified, pretested, and post-tested. The design was visualized as:

Test	Treatment	Post-test
0	X_1	0
	X_1	0
0	X_2	0
	X_2	0
0	X_3	0
	X_3	0
0	X_4	0
	X_4	0

with only part of each of the four groups pretested, and as many as remained post-tested. (Initial pretest selection was on a random basis.) Additional students who were not pretested, but who received the various treatments, were post-tested. All possible efforts, under the constraints of ongoing classroom activities, were made to ensure internal and external validity.

SWCEL test administrators were able to test more students each day, and consequently spent less time on each test administration. The result was a significant influence between groups of testers, even when testing children from the same schools.

This was remedied on the post-test by having all students from Group I, III, and IV tested by SWCEL personnel. (Post-testing in Group II was done by Service Center personnel. Since 16 students were pre- and post-tested, they were, therefore, dropped from further analysis.)

Other variables, which were not controlled but must be noted, included variation in urban-rural status between various groups. Students were tested from communities having a population range of 100 to 150,000.

Assessment Instrument

The assessment instrument was the Michael Test of Oral English Production (MTOEP).

SWCEL's use of the Michael Test. Because there was nothing else available at the time (and nothing new has been developed at this date)

the Laboratory used the Michael Test during the spring of 1969 to evaluate the ROCK project for the Region One Education Service Center.

The Michael Test was never thought to be a general competency test covering the entire range of possible English syntactic constructions. Its utility as a vehicle for assessing performance in spoken English is only within the range of verbal behavior as encompassed by the scope of the H-200 Oral Language Program.

Lois Michael's claims for the test:

> There are no standardized test materials to measure the oral English proficiency of primary age children. Fortunately, the lessons in Project H-200 list specifically stated instructional objectives that define both terminal behavior and content. Those objectives served as the basis for developing the test . . . (page 4, Summary Report by the late Lois Michael).
>
> As a measure of content validity the test was administered to 30 first graders who are native speakers of English . . . selected randomly from two first grade classes in California in one of the districts from which part of the control came. The native speakers, regardless of low or high IQ, performed uniformly well on the test. All had a score of 99-100 percent of the total . . . possible points. (page 5, Summary Report by the late Lois Michael.)

Experience with the test indicates the following. Lois Michael's claim that the test content is valid for native English speakers, on the basis of 99-100 percent test scores having been obtained for these speakers, has not been substantiated by data. Whereas the total possible test score is 224 points, the highest score on file is 218. It was found that children who score more than 200 are very good speakers--i.e., those who have internalized the syntactic rules for producing most well formed sentences.

The primary reason why very good speakers may score poorly on the test is because 20 percent of it consists of items requiring the generation of questions. SWCEL testing experience has shown that most of the students in the three to nine age range have not as yet acquired the syntactic competence to respond appropriately to the test stimulus for the Michael Test items designed to elicit questions. Children are expected to produce questions in response to these cues: "Ask me if . . . ," "Ask me how many . . . ," "Ask me who . . . ," etc. Responses received to this type of stimulus include:

1. Silence.
2. An attempt to ask a question but in some unacceptable form, for example, omitting the auxiliary or failing to switch the order of subject and verb, although the response will have the appropriate intonation for a question, e.g., "He has a pencil?" for "Does he have a pencil?"
3. An answer (usually logical) to the administrator's statement.
4. A repetition of the tester's statement.

Performance data. Michael Test scores will indicate, in a general way, that it is possible to group speakers in certain categories. The most general kind of grouping would be a threefold division:

GROUP I Scores between 0-100. This group includes speakers with little or no knowledge of English.

Any test points would come largely from the pronunciation category where a point can be scored by repeating the test item correctly, from the vocabulary section by identifying the test item correctly, or from the ubiquitous communication category.

GROUP II Scores between 101-200. This includes a very wide range of substandard speakers.

GROUP III Scores between 201-224. Those in this range are "good" speakers whose command of the language obviates any need for ESL materials (Level I as defined by ROCK materials). Experience indicates that some of the students who score in this range may need additional ESL, but greater validation is necessary.

A closer analysis of scores in the 101-200 range suggests the following subdivisions:

GROUP A Scores between 101-130. Speakers in this group have difficulty comprehending many of the test items. Attempts to elicit types of construction frequently will be met with silence or a repetition of the test item. However, Group A is sufficiently in control of the language to communicate via ill-formed syntactic constructions.

GROUP B Scores between 131-160. Speakers in this group both comprehend and respond to test items better than Group A. However, they often do not respond to test items without use of one of the "branching" procedures. Although they tend to use a large number of ill-formed constructions (especially speakers at the lower end of this range), these deviant forms will alternate with their well-formed counterparts. Their language facility could be described as being in a state of flux.

GROUP C Scores between 161 and 200. These speakers both comprehend and respond to the test items. They have internalized the rules for most well-formed constructions, and their syntactic lapses are of a relatively minor kind. These lapses are, however, of the type that will probably persist into adult speech marking them as slightly deviant by middle class standards. Examples include: "I ain't . . ." "Did you got . . ."

Conclusions. The Michael Test has been used extensively (on more than 2,000 children) to reach the foregoing conclusions concerning its utility in ascertaining certain levels of linguistic performance. However, despite certain claims Lois Michael made regarding the test's validity and reliability, certain factors need attention over a given period of time. For example:

1. It has not been determined whether the size of gains in points is comparable in different score ranges.
2. There are no norms available as to how well children should do after "X" number of lessons, or how much they should gain, but this needs to be established.

The lack of other criteria by which to establish comparisons has, therefore, led to the selection of the MTOEP as the assessment measure for this evaluation.

Limitations of the Assessment Plan

The greatest limitation is in itself the reason for the program--individual variation and differences. Students from a migrant population are highly mobile. Many students continue to enroll throughout the school year, and accordingly, dropout rate and attendance for this are control problems. This is indicated by noting the variation in length of enrollment of students at pretest time from 0-93 days.

Teachers also are very individualistic. Not all participating teachers will teach the lessons as scheduled. This is indicated by noting the lesson variation of classes using the materials as long-range from lesson 0 to lesson 44, i.e., some teachers were on lesson 44 when the pretest was given while others had failed to start the materials.

Such variations continued through post-test time. Some students had had high attendance, others low. Some teachers had progressed through many lessons, others few.

The testers also are variables. Two groups of testers were involved, one from the regional service center staff and a second from SWCEL.

Assessment Summary

Considering all the variables listed above, and specifically the differing linguistic competence among children, two assessment techniques of performance criteria and experimental versus control were employed. The Michael Test of Oral English Production was chosen as the evaluative instrument as it is the only test purporting to measure this specific phenomena.

GROUP INFORMATION

Migrant Population

The children involved in this investigation were five-year old Mexican American native speakers of Spanish with little or no knowledge of English. They were from the extremely mobile population of migrant farm workers who follow the crops from the Lower Rio Grande Valley north throughout the United States. As a result, the child is in school for brief periods of time, interrupted by frequent moves throughout the year. This mobility contributes, also, to the culture-language cohesion of the group which remains generally outside the mainstream of American English-speaking culture complicating the schools' language-socialization process.

Tested Groups

Group I was composed of children living in Region One's seven counties along the Lower Rio Grande River, winter home base for a majority of the migrant workers, an area maintaining close ties with the root culture across the Rio Grande where the child is immersed in a Spanish speaking society both at home and at school. Migrant pre-school teachers in Region One were largely native speakers of Spanish, most of them limited in their ability to produce oral English. At the time of pretest the range of number of days in school for Region One children was from 0 to 94 with a mean of 33. Before beginning the ROCK program teachers in groups of 25 to 50 were given 16 hours of

training in oral language techniques and specific use of ROCK materials. During the year very limited classroom supervision was given by Region One consultants, with two consultants serving 122 teachers. Thirty percent of classrooms in this group did not have an audio flashcard machine, and 50 percent did not have the sound filmstrip projectors with headphones, both required for full implementation of the ROCK program.

Group II contained five-year olds of the same migrant Spanish speaking population, but they were tested in schools of Regions Two and Twenty where many of them have their permanent homes. Teachers in this group received the same training from Region One's staff but no classroom supervision. ROCK materials were used.

Group III was composed of five-year old migrant children who were tested in the Texas Panhandle, the point farthest away from the Mexican border, in a location where they are surrounded by an English speaking population, and where teachers are usually competent in the production of English. These teachers received neither training nor supervision from Region One. ROCK materials were used.

Group IV consisted of three school districts in Region One (location of Group I) and a district in Region Twenty (location of Group II). The children were five-year old Mexican American Spanish speakers. They were not receiving instruction with ROCK materials. Instead, their teachers were using a variety, including the Bereiter and Englemann DISTAR Language Program, local district-prepared guides

for English language teaching of the categorized-vocabulary type, or traditional pre-basal reading readiness procedures.

(See Tables I and II on the following pages.)

Summary. In essence Groups I, II, and III used ROCK materials with a range of zero to 16 hours of teacher training. Those in Group IV received no teacher training. Only those in Group I received any supervision by consultants. While the youngsters were all Mexican American, Spanish speaking migrant children, their environmental surroundings ranged from those practically 100 percent Spanish speaking, to those 100 percent English speaking everywhere except at home. These factors should be kept in mind when interpretations of the test data are reviewed.

TABLE I

Group I Test Data

District	Population	No. of Schools	No. of Classrooms	ROCK Lessons		Test Dates		Test Scores	
				Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test
Edinburg	18,706	1	1	29	68-75	1/ 9/69	4/ 8/69	100	159
Harlingen	41,207	2	3	4 15 10		12/16/68	4/ 8/69	69	88
Laredo	152,500	1	5	29 23	61-64 61-64 63 62	1/17/69	4/ 8/69	96	
La Sara	200	1	1	9	63	1/10/69	4/ 9/69	140	150

Group I Test Data, Continued

District	Population	No. of Schools	No. of Classrooms	ROCK Lessons Pre-test	ROCK Lessons Post-test	Test Dates Pre-test	Test Dates Post-test	Test Scores Pre-test	Test Scores Post-test
Mission	14,081	4	7	11 4 24 34 23 35 0 14	90 71 72 79 43 50 45 48 49	1/13/69	4/ 7/69	144	
Rio Grande City	5,835	2	8	11 13 15 10	22 15 23 23 25 32 28	1/20/69	4/ 8/69	128	150
Donna	7,522				14	1/22/69	4/11/69		
Rio Hondo	1,344	1	1	11	82		4/10/69	134	157
San Benito	16,422	1	3	29		1/24/69	4/ 8/69	143	76

Group I Test Data, Continued

District	Population	No. of Schools	No. of Classrooms	ROCK Lessons Pre-test	ROCK Lessons Post-test	Test Dates Pre-test	Test Dates Post-test	Test Scores Pre-test	Test Scores Post-test
San Juan	4,371	1	4	3 0 1 5		12/17/68	4/10/69	93	168
Weslaco	15,649	1	2	12 7	36 71 46 58 50	1/15/69	4/10/69	103	156
Pharr	14,106	2	8	16 17 14 11	13 9 3 10 42 47 44 47 66	1/ 8/69	4/ 9/69	131	153
Edcouch	2,814					1/10/69	4/11/69		

TABLE I

Group II Test Data

District	Population	No. of Schools	No. of Classrooms	ROCK Lessons		Test Dates		Test Scores	
				Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test
Carrizo Springs	5,699	2		17		1/ 9/69		139	
Crystal City	9,101	1	4	36 34		1/23/69		129	
Poteet	2,811	1	1	13		1/30/69		141	
Robstown	10,266	2	3	23 12 0		12/17/68		113	171
Sinton	6,008			0		1/24/69			
Eagle Pass	12,094	1	1		Present	1/14/69		83	

TABLE I

Group III Test Data

District	Population	No. of Schools	No. of Classrooms	ROCK Lessons		Test Dates		Test Scores	
				Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test
Brownfield	10,286	3	3	0		1/16/69	4/16/69	152	153
Dimmitt	2,935	1	2	0		1/20/69	4/17/69	143	123
Hereford	7,652	5	15	0		1/20/69	4/16/69	126	152
La Mesa	12,438	1	2	0		1/15/69	4/16/69	127	150

TABLE I

Group IV Test Data

District	Population	No. of Schools	No. of Classrooms	ROCK Lessons		Test Dates		Test Scores	
				Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test
Mercedes	10,943	4	7	OLP Present		1/30/69	4/10/69	161	155
Monte Alto	150	1	1	0		1/29/69	4/10/69		166
Progreso	900	1	1	0		1/29/69	4/ 8/69	123	137
Uvalde	10,293	3	10	0		2/ 4/69	4/16/69	145	159

TABLE II

Materials, Training, and Supervision
Received by Test Groups

	Group I	Group II	Group III	Group IV
Materials	ROCK	ROCK	ROCK	Non-ROCK
Training	16 hours	16 hours	None	NA
Supervision	Two Consultants and 122 Teachers	None	None	NA

PROCEDURES FOR GATHERING DATA

Training of Testers

Pupils in the study were tested by examiners trained in one two-day institute and one three-day institute at the Southwestern Cooperative Educational Laboratory in Albuquerque on the Michael Test of Oral English Production. All testers worked under observation with bilingual children and practiced test administration before actually administering the test. (See Appendix A for training program agenda.)

The Michael Test, developed by the late Lois Michael, a graduate linguistics student at UCLA, is actually an achievement test concerning materials taught during the first year of H-200 lessons.

The SWCEL staff, along with those initially involved with Miss Michael and the test (Dr. John Otis and Dr. Robert Landen), made certain adjustments to the basic test form to facilitate scoring and testing and to allow more standardization of test administration.

Reliability of Scores

The test is given individually, recorded on a tape, and scored later. Scoring and administration must be standardized. Three scorers were trained by institute training in administration and scoring and after a minimum of one week of actual test administration. They had to reach a criterion of 95 percent accuracy with five actual test tapes. Periodic checks were made to ensure reliability.

Administration of the Test

The pretests were given by 10 trained test administrators beginning December 16, 1968, and ending February 4, 1969. The test tapes, scored by three persons, were graded between January 13, 1969, and April 1, 1969.

Post-tests were given by eight persons from April 8, 1969, to April 16, 1969, and scoring, done by the same three scorers, was completed by May of 1969.

Pretests were given to 414 children; 283 of these children were post-tested with an additional 112 children post-tested.

The difference between the number pre- and post-tested is because of children following the crops with their parents.

Limitations

Since a migrant population was being dealt with the length of the children's enrollment varies. Many of them attend school sporadically, and enroll as late in the year as January.

Although post-testing covered only eight days, the same groups were tested first each time, and all scoring was done by the same three scorers. Pretesting was done over a period of seven weeks. Teachers covered different numbers of lessons, used lessons differently, had differing amounts of classroom help, and different types of equipment.

Although it is not impossible to measure teacher attitude at the time of testing, such a factor does tend to have some effect on the children's reaction to the test administrator and to the test.

If the teacher is open, friendly, and enthusiastic, the children feel they are special. If the teacher is defensive, however, the children tend to be more wary and uncommunicative during the test.

The degree to which responses were affected by where the test was given (nurse's clinic, principal's office, etc.), is hard to determine, but was considered random across all populations.

RESULTS AND INTERPRETATION OF DATA

Analysis

Students assessed for the evaluation either received the pretest only (N=127), received the post-test only (N=112), or received the pretest and the post-test (N=283). Of the 414 students pretested, only 283 remained in the sample at post-test time. Three hundred ninety-five students were post-tested; 112 students received the post-test but not the pretest. A total of 522 students participated in the sample, and 805 tests were administered in the Edinburg area.

There was no significant difference at the .05 level of confidence between those students receiving the post-test only and those who received the pretest and post-test, indicating that the students did not learn enough about the test from the pretesting situation to make any significant differences on post-test scores. (See Table III and its related graph.) When examining the students who were pretested and post-tested it was found that Group I contained 122 students, Group II contained 16 students, Group III contained 82 students, and Group IV contained 73 students. Group II was discarded from further analysis because of comparatively small sample size. This reduced the sample size from 283 to 267 for the final analysis.

Of major concern was the differences between the three remaining groups on the Michael Test. The test was analyzed by examining two of the scores, the partial score on Structure, and the Total Score. The Michael Test contains four subparts: Communication, Structure, Vocabulary, and Pronunciation. The algebraic sum of these four scores

TABLE III

Pretest and Post-test versus Post-test Only
on Total Michael Test Score

Analysis of Variance

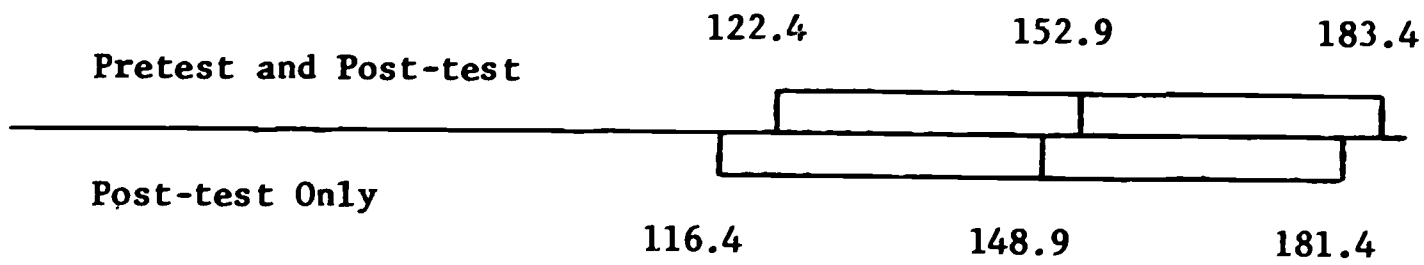
Source	df	ss	ms	F
Treatment	1	1289	1289	1.34
Error	393	378900	964	
Total	394	380189		

General Statistics

	Number	Mean	Standard Deviation
Pretest and Post-test	283	152.9	30.5
Post-test	112	148.9	32.5

TABLE III--GRAPH

Range of Scores of
Two-thirds of Pupils



yields a Total Score. The Structure section was identified by the evaluators and program developers as the most important part of the test proper. This score is probably most indicative of the student's progress in the Oral Language Program being evaluated (ROCK). The Total Score is, however, a reflection of the student's overall progress in the four areas previously mentioned.

Evaluation of the students in each of the three remaining groups is considered in four ways. First, differences on the pretest; secondly, differences at the post-test; thirdly, differences between the groups when post-test scores are adjusted for pretest scores; and fourthly, differences between the groups on gain scores.

1. Pretest scores. The difference between the three groups was not significant at the .05 level of confidence on the Structure subscore at pretest time. Both experimental groups did score lower than the control group (Group IV). (This is summarized on Table IV.) Analysis of variance of the Total Score at pretest time was significant at the .01 level of confidence, with Group I being significantly lower than Groups III or IV, and with more significant difference between Groups III and IV. (This is summarized on Table V.)

2. Analysis of post-test data. There were no significant differences between the three groups on the Structure subscore or on the Total Score at post-test time. This indicates that differences which existed at pretest time

TABLE IV

Pretest Structure Scores
on Michael Test

Analysis of Variance

Source	df	ss	ms	F
Treatment	2	842	421	2.18
Error	264	51010	193	
Total	266	51852		

General Statistics

	Number	Mean	Standard Deviation
Group I	112	19.21	15.02
Group III	82	20.60	11.44
Group IV	73	23.56	14.34

TABLE V

Pretest Total Scores
on Michael Test

Analysis of Variance

Source	df	ss	ms	F
Treatment	2	23100	11550	8.15
Error	264	374100	1417	
Total	266	397200		

General Statistics

	Number	Mean	Standard Deviation
Group I	112	118.08	42.18
Group III	82	135.28	27.59
Group IV	73	138.45	39.25

had been removed, i.e., Group I, which was significantly lower, was now not significantly different from either of the two remaining groups. (This is summarized on Table VI.)

3. Analysis of adjusted post-test scores. By using the statistical treatment of analysis of covariance, a significant difference was found between the three groups on the Structure post-test scores. Although there was no significant difference between the experimental groups, there is a significant difference between the experimental groups (Group I, Group III) and the control group (Group IV), with the adjusted Structure means of the experimental groups being significantly greater. (This is summarized on Table VII and its related graph.)

A similar analysis of total post-test scores indicated a significant difference between experimental and control groups, with Groups I and III significantly higher than Group IV. Had all the groups been equal at pretest time, it is statistically estimated that Group I would have scored 156.7 average, with Group III averaging 152.8, and Group IV averaging 142.7. These, however, are statistical "ifs." (This data is summarized in Table VIII and its related graph.)

4. Analysis of gain scores. If the pretest score is subtracted from the post-test score, a measure of student gain is given for that test. Every student participating

TABLE VI

Post-test Total Scores
on Michael Test

Analysis of Variance

Source	df	ss	ms	F
Treatment	2	3434	1717	1.84
Error	264	245900	931	
Total	266	249334		

General Statistics

	Number	Mean	Standard Deviation
Group I	112	149.48	32.52
Group III	82	157.05	22.68
Group IV	73	148.99	34.18

TABLE VII

Adjusted Post-test Structure Scores
on Michael Test

Analysis of Covariance

Source	df	ss	ms	F
Treatment	2	1413	707	10.79
Error	263	17220	65	
Total	265	18633		

General Statistics

	Number	Adjusted Mean
Group I	112	29.85
Group III	82	30.74
Group IV	73	25.10

TABLE VII--GRAPH

Comparison of Groups by Groups on Pre- and Post-test Structure Scores

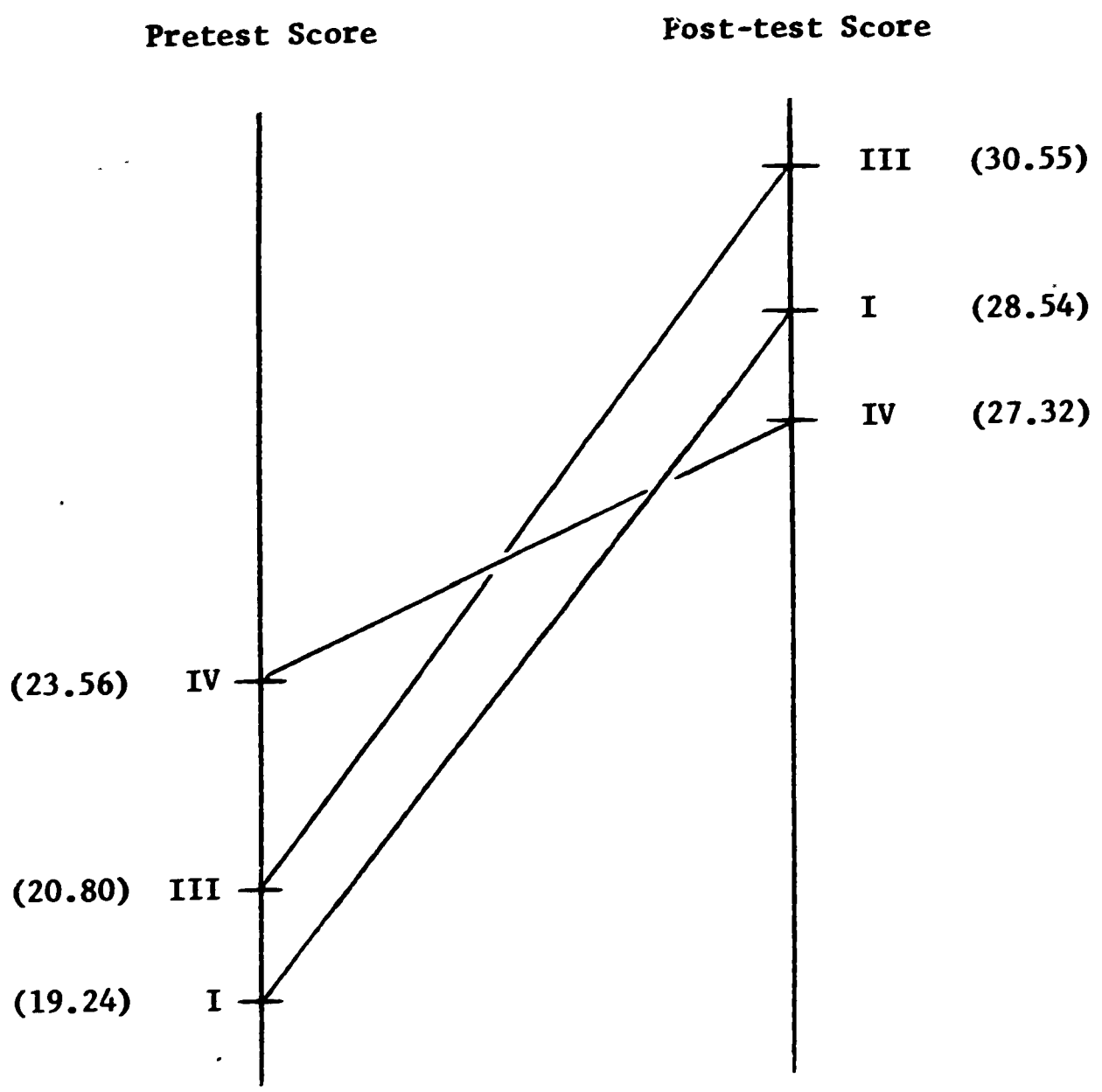


TABLE VIII

Adjusted Post-test Total Scores on Michael Test

Analysis of Covariance

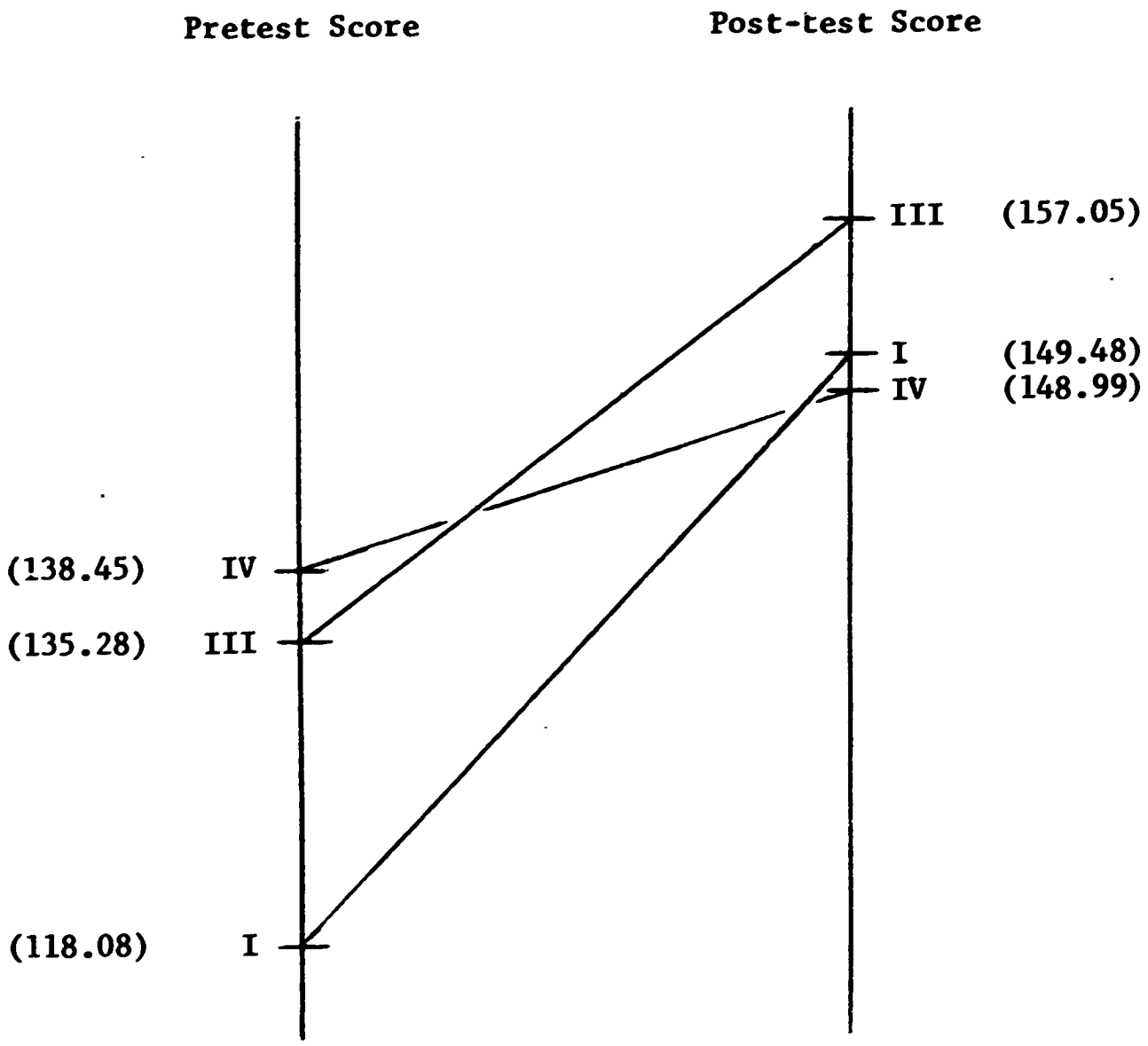
Source	df	ss	ms	F
Treatment	2	8466	4233	13.62
Error	263	81760	311	
Total	265	90226		

General Statistics

	Number	Adjusted Mean
Group I	112	156.67
Group III	82	152.84
Group IV	73	142.68

TABLE VIII--GRAPH

Comparison of Groups by Groups on
Pre- and Post-test Total Scores



in the pretest and the post-test has a gain score. A natural assumption would be that students tested about three to four months after initial testing would score higher on a language production test administered four months later if they had been practicing over the four-month interval on material which was to be assessed. This assumes exactly equal testing conditions and an identical mental frame of reference for the student during both testing sessions. Some students scored lower on the post-test than on the pretest.

An analysis of variance of the gain scores indicated that Groups I and III, the experimental groups, gained significantly more on the subscore of Structure than did the control group (Group IV). (This data is summarized in Table IX.)

An analysis of variance of the gain scores from the pretest to post-test of the Total Score on the Michael Test indicated a significant difference between all groups. Group I students gained the most with an average of 31.4 points from pretest to post-test. Group III had an average gain of 21.8 points from pretest to post-test, and Group IV students had an average gain of only 10.5 points. This indicated a significant difference at the .01 level of confidence between Group I and Group IV, and a difference at the .05 level of confidence between

TABLE IX

Analysis for Independent Variable Region and
Dependent Variable Structure Scores on Michael Test

Analysis of Variance

Source	df	ss	ms	F
Treatment	2	1824	912	12.65
Error	264	19020	72	
Total	266	20844		

General Statistics

	Number	Mean	Standard Deviation	High Score	Low Score
Group I	112	9.32	9.63	35.0	-10.0
Group III	82	9.95	8.94	32.0	-16.0
Group IV	73	3.75	5.37	16.0	- 9.0

Groups I and III, and Groups III and IV. (These data are summarized in Table X.) Students in Group I had significantly higher rates of learning than students in either of the other two groups as assessed by the Michael Test.

An interesting secondary finding was analysis of scores between boys and girls. No significant differences were found on pretest scores, on post-test scores, on adjusted post-test scores, or gain scores for either Structure or total test scores. Perhaps the data can best be summarized by a table and graph which indicate differences between pre-test and post-test scores for the three groups analyzed. (See Table XI and its related graph.)

Summary. The data presented are raw scores, and in no way adjust for unequal groups to begin with; nor do they take into account that groups which scored low on the pretest should find gain to be easier. These data directly indicate that students who received the ROCK materials started with average scores lower than the control group, but in only three and a half months had increased in language proficiency so that their scores at the post-test time were equal to or exceeded those of the control group.

TABLE X

Analysis for Independent Variable Region and
Dependent Variable Total Scores on Michael Test

Analysis of Variance

Source	df	ss	ms	F
Treatment	2	19350	9677	20.53
Error	264	124400	471	
Total	266	143750		

General Statistics

	Number	Mean	Standard Deviation	High Score	Low Score
Group I	112	31.40	27.44	111.0	-15.0
Group III	82	21,77	17.21	61.0	-10.0
Group IV	73	10.53	14.70	63.0	-20.0

TABLE XI

Adjusted Total Scores by Sex
on Michael Test

Analysis of Covariance

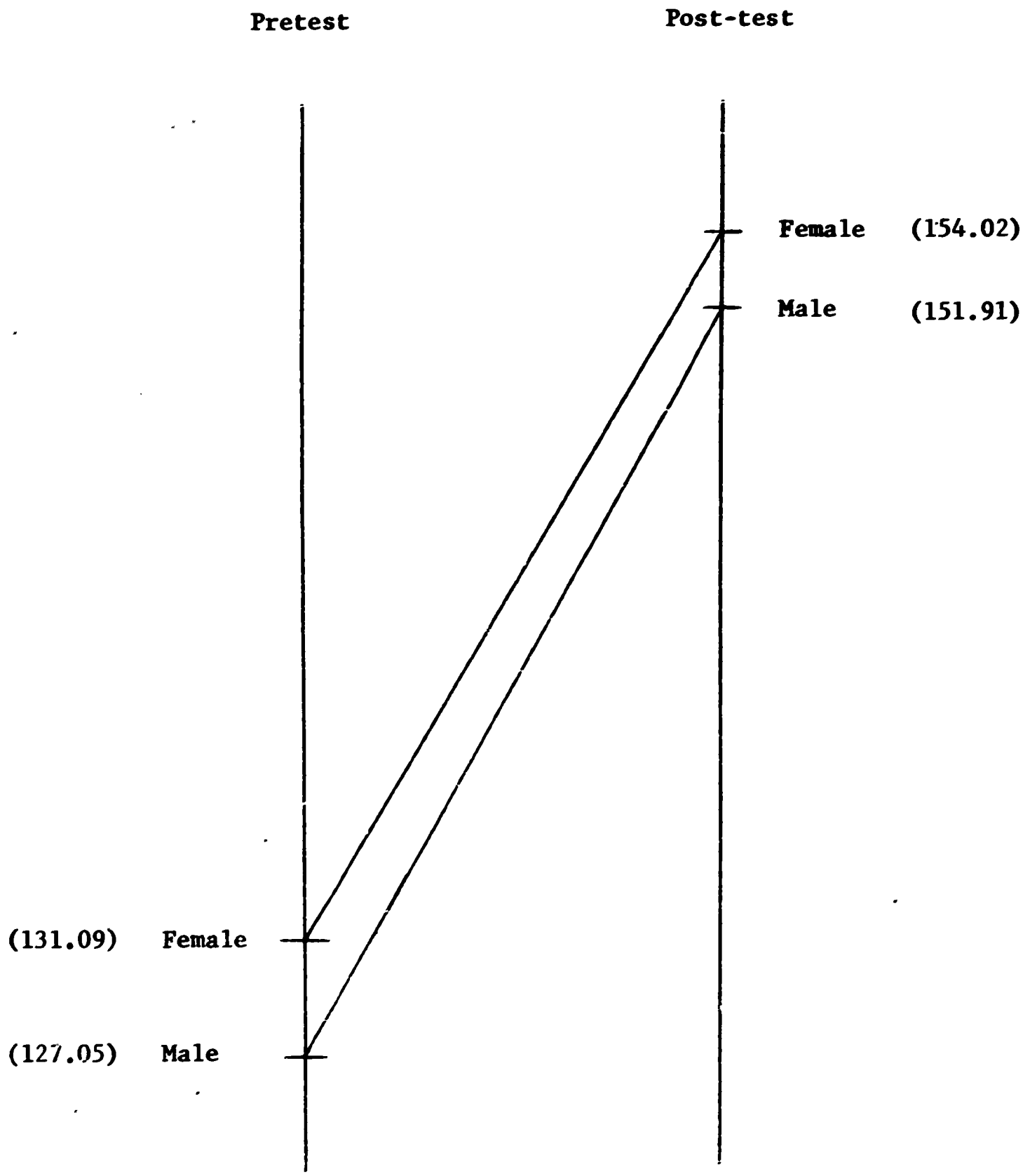
Source	df	ss	ms	F
Treatment	1	15	15	.041
Error	280	99680	356	
Total	218	99695		

General Statistics

	Number	Adjusted Mean
Male	150	153.11
Female	133	152.66

TABLE XI--GRAPH

Comparison of Males and Females on Pre- and Post-test Total Scores



SUMMARY AND RECOMMENDATIONS

Many of the more than five million Mexican Americans--our second largest minority group--have less than an equal educational opportunity; the needs of the migrant and the concomitant educational problems are especially acute. These include such factors as transiency, poverty, cultural isolation, and language barriers. Coupled with this is the realization that all too often there are less than adequate materials and less than adequately trained teachers, and we are soon in a nearly overwhelming situation if careful attention is not directed at crucial elements.

The Southwestern Cooperative Educational Laboratory believes the Region One Education Service Center has accomplished the following:

1. Focused on an extremely significant educational problem--that of migrant education at the preschool level.
2. Avoided the temptation to "reinvent the wheel" by adapting, expanding, and improving previously prepared educational materials, specifically those which went through the design stages at UCLA under the direction of Dr. Robert Wilson and Eddie Hansen. (Dr. Wilson is with the English Department of the University of California at Los Angeles, and Mr. Hansen is with the State Department of Public Instruction in California.)
3. Identified exceptionally competent program personnel, such as Al Ramirez, who provide the necessary balance between scholarly expertise and concrete understanding of school problems.

4. Initiated the beginning steps for helping to resolve the educational problems of the target population.

While exceptional progress has been made, particularly considering the antecedent obstacles that have plagued progress for years, much remains to be done and the Southwestern Cooperative Educational Laboratory would recommend the following:

1. Establishing reliable accountability and responsibility criteria. Apparently some teachers failed to use the materials as their use was intended. If fiscal and human resources are to be increased for development of relevant curriculum materials and the extension of energies and services in pre- and in-service training, one might reasonably expect materials to be used close to the design specifications established. Consideration should be given to the establishment of quality control procedures for this purpose.
2. Continuing the development of materials which include the development of criterion tests, recycling materials for students who have not achieved acceptable performance standards, and designing other types of supplementary materials for teacher use.
3. Continuing the assessment and evaluation activities that have been started. Longitudinal evaluation is seldom done in American education but lasting effects cannot be expected from short-term episodes when certain follow-up activities

of articulation and continuity are not maintained. The need for longitudinal evaluation is obvious.

4. Improving the testing instruments now available. While not perfect, the instruments are the best available, and it is recommended that present instruments be more completely developed rather than beginning a new process.
   a. The logistics of administering the test need to be refined. For example, the test should be designed so that teachers can easily, quickly, and accurately administer and score their tests and make judgments about students' needs.
   b. The test needs to be shortened and revalidated. Perhaps the section on Pronunciation should be eliminated from the format, although careful professional thought should be given to this linguistic question.

While the data gathered neither particularly confirm nor support the following points, it is felt the factors should be mentioned and given a place in the report as the ideas cropped up frequently in discussions with many of the personnel who participated in the program.

1. Considerable attention should be directed at a specified teacher training program wherein teachers and aides receive comparative training. Some efforts have been made, but limited staff and fiscal resources prohibited the extensive type of desirable teacher training activities. Only teachers who have been trained in the use of the ROCK methods should

be permitted to use them. Hopefully, this will avoid the problems many educators have observed when such nationally designed curricular materials as BSCS, ITA, and SMSG were installed in classrooms. While the materials may have been relatively good, the teachers did not have the appropriate instructional strategies--perhaps a gross mistake.

2. Once the technical skills of the teacher and aide have been established some procedure, such as periodic in-service training programs, is needed so such skills can be maintained. This probably will necessitate two factors:

   a. Determining teacher needs and designing in-service packages accordingly.

   b. Prolonging in-service activities; ultimately this will probably require the use of coordination and additional consultants to give in-the-classroom support. To expect two consultants to adequately monitor and provide in-the-classroom assistance at the ratio of 60 teachers to one consultant obviously is unrealistic, particularly during the embryonic stages of installation and service testing of materials.

The recommendations call for additional resources--human and fiscal. It is the Southwestern Cooperative Educational Laboratory evaluation team's opinion, however, that the limited resources invested at the Region One Center have been expended wisely to date--even with the known limitations spelled out throughout the report--and it is

recommended that the Texas Education Agency should continue as well as expand the prototype work under way.

Many of the limitations outlined reflect the problems of action-oriented educational research, which in one opinion is more likely to produce the desired results than the sheltered, controlled research typical of education--even though arguments for both activities can be made.

Finally, these statements seem in order:

The results achieved thus far reflect the soundness of the concepts underlying the research and development activities at Region One as well as the energy and capability of staffs working under less than ideal conditions. To realize full potential of the ESC agency, measures need to be taken promptly to build its essential institutional stability and continuity, and to change conditions which place a drag on its effectiveness. For example:

1. Assure operation and basic funding for periods of three to five years.
2. Provide for orderly and significant increases in the support level of the organization.
3. Establish the conditions which are necessary to accommodate accountability and freedom from bureaucratic constraints which hamper creativity and productivity.

We applaud the work that has been done by those at the Region One Service Center and we hope, by way of this report, that we have detailed some of the work which remains.

APPENDIX A

TRAINING PROGRAM FOR MICHAEL TESTERS

SWCEL EDINBURG TEST INSTITUTE

Monday, December 2, 1968

8:30 -- 8:45 A.M.	Introductions
8:45 -- 10:00	Discussion of standardized test procedures
10:00 -- 10:15	Coffee Break
10:15 -- 10:45	Tape recorder demonstration
10:45 -- 11:15	Videotape demonstration (show Randall tape)
11:15 -- 12:00	General Discussion
12:00 -- 1:00 P.M.	Lunch
1:00 -- 3:00	Microtesting and videotaping
3:00 -- 4:30	Discussion and observing tapes

Tuesday, December 3, 1968

8:30 -- 12:00	Further testing experience
12:00 -- 1:00	Lunch
1:00 -- 3:00	Microtesting and final videotaping
3:00 -- 3:30	Closing discussion

SWCEL MICHAEL TEST TRAINING SESSION

(Upstairs Conference Room)

Tuesday, January 7, 1969

8:15 -- 8:30 A.M.	Tour of Laboratory
8:30 -- 8:45	Introductions (Dr. Liberty and Dr. Reeback) Upstairs Conference Room
8:45 -- 9:00	Discussion of SWCEL objectives (Dr. Seaberg)
9:00 -- 9:30	Discussion of standardized test procedures (OLP relationship with Michael Test)
9:30 -- 10:30	Videotape of Michael administration
10:30 -- 10:45	Coffee Break
10:45 -- 11:00	Tape recorder demonstration
11:00 -- 12:00	Paired practice
12:00 -- 1:00 P.M.	Lunch
1:00 -- 1:30	Branching and other fine points of administration
1:30 -- 2:30	Paired practice
2:30 -- 2:45	Coffee Break
2:45 -- 4:00	Discussion of scoring Phonetics of test items

Wednesday, January 8, 1969

8:45 -- 12:00	. . . . .	Testing experience with children
12:00 -- 1:00	. . . . .	Lunch
1:00 -- 2:30	. . . . .	Scoring conference and practice
2:30 -- 2:45	. . . . .	Coffee Break
2:45 -- 4:00	. . . . .	Review of scoring Preview of criterion tests

Thursday, January 9, 1969

8:45 -- 12:00	. . . . .	Testing experience with children
12:00 -- 1:00	. . . . .	Lunch
1:00 -- 1:30	. . . . .	Review test administration
1:30 -- 4:00	. . . . .	Scoring conference and practice